Modelling Spatiotemporal Fluctuations of Consumer Demand in the UK Grocery Sector and their Impact on Retailers Store Sales

By

Thomas Bryan Platon Waddington

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

School of Geography

August 2017

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly authored publications has been included.

The contribution of the candidate and the other authors to this work has been explicitly indicated below.

The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Chapters 4 and 5 draw on analysis and findings which has been reported in the following publication.

Waddington, T.B.P., Clarke, G.P, Clarke, M. and Newing, A. 2017 Open all hours: Spatiotemporal fluctuations in UK grocery stores and catchment area demand. International Journal of Retailing, Distribution and Consumer Research, 1-26.

The first named author has carried out the research within the paper. The manuscript was prepared by the first named author, with the input of the co-authors being advisory and editorial.

Acknowledgements

I would like to express my gratitude to all those individuals, either directly or indirectly, who have helped me at any stage during my thesis, be it academic or personal, their support and continued presence has been crucial.

I would like to thank the amazing support staff at the University of Leeds whom have been there at various stages of my theses, offering support and advice and making the whole process much more achievable and clear. Specifically, the administrative help of Jacqui Manton and the statistical help of Rachel Homer have been unfaltering and greatly appreciated. Their advice and expertise has certainly helped me manage my PhD and enabled the analyses presented here.

I would like to thank the encouraging and ongoing support of my supervisors. Graham and Martin have been supportive throughout my thesis and given a lot a freedom to learn and explore, allowing me to take my research down the route that I wanted. Their combined experience and advice has allowed me to develop as a researcher and ensured that I complete my thesis on time and with a broad range of skills and knowledge that I now carry on to the next stage of my life. Their relaxed approach has worked really well and even with many distractions there have been very few times when I felt that the process was becoming too much thanks to their guidance.

I would also like to acknowledge the contribution of our data partner and staff, whom without their generous support, time, and almost unlimited response to data requests, this research would not have been possible. I would like to thank them for the freedom to explore in detail their data, which enabled me to identify interesting novel research agenda and through a continued friendship to tailor my research ensuring applied and ongoing positive implications in the commercial sector.

A special mention and huge appreciation is for Andy Newing. His selfless assistance and support has been hugely influential in the development of my thesis and of myself and without his guidance, huge depth of expertise and friendship, this would have been a very different experience and stage in my life making it all the more difficult.

Similarly, I'd like to thank all my friends at the university, PhD students and those from the Ski society that have supported me academically and socially making the tough and stressful times far more enjoyable.

Finally I would like to thank my family and my girlfriend Tara, without your support this would not have been possible. You have all been there for me without hesitation, helping me financially, emotionally and socially, making me laugh and ensuring that I didn't let the stress take over and I appreciate that so much. I hope that I haven't been too much of a burden over the years and I promise that I will give as much in return as you have given me.

Abstract

Retail location planning in the grocery sector is used to aid location-based decision-making; addressing issues such as, store performance, store planning or for estimating demand. Location planning teams employ a range of sophisticated modelling techniques; this thesis draws on Spatial Interaction Models (SIMs). Spatiotemporal components have received limited attention in relation to retail location planning and this is addressed within this thesis. It is argued that spatiotemporal components of consumer demand fluctuate considerably within catchment areas, which has a direct impact of store sales, and that models, which have accounted for specific aspects of spatiotemporal demand result in better and more representative store revenue estimations.

This thesis demonstrates the importance of spatiotemporal demand through an analysis of store level diurnal sales patterns which are related to observed consumer shopping habits, providing novel insight and improving our understanding of supply and demand-side spatiotemporal components. The insight is used to generate a series of spatiotemporal demand layers, which are used in conjunction with a (SIM) to generate spatiotemporally informed store level revenue estimations.

Rare access to temporal EPOS transaction records at a store level and substantial loyalty card data provided by a major player in the UK grocery market presents novel opportunities for analysis. The data enables this thesis to generate insight into temporal fluctuations of store sales and the demand side drivers behind them, as well as, the geographies of consumer demand. The findings demonstrate evidence of distinct clusters of diurnal sales profiles in stores, which appear to be directly related to specific spatiotemporal demand types, and is used in conjunction with a SIM. The analysis shows that through adding spatiotemporal demand layers, demand estimates within catchment areas are far more representative and can lead to improved revenue estimations.

Table of Contents

<u>**List of tables**</u>

List of Figures

List of Abbreviations

PPG – Planning policy guidance

ATD – Average trip distance

MPS – Multi-purpose shopping

OA – Output area

LSOA – Lower super output area

MSOA – Middle super output area

LAD – Local authority district

MPTM – Multi-purpose trip making

SIM – Spatial interaction model

WP – Workplace

WPZ – Workplace zone

OAC – Output area classification

ONS – Office of National Statistics

RPM – Retail price maintenance

FSA – Food Standards Agency

GIS – Geography information system

MAUP – Modifiable areal unit problem

GOF – Goodness of fit

COWZ – Classification of Workplace Zones

SQFT – Square foot

TI – Trade intensity (a measure of store performance using £/Sqft of floorspace)

MP – Multi-purpose

GRO – General register office

HESA – Higher education statistics agency

Y + H – Yorkshire and the Humber

Chapter 1 – Introduction

1.1 Research outline and context

The UK grocery sector has undergone considerable change over the last few decades, incorporating periods of significant growth and decline in market size, performance and in retail operation (Birkin et al., 2002, Birkin et al., 2017, Hood et al., 2015, Keh and Park, 1997, Wrigley, 1994). Often, the processes of decision-making that influences the development, closure or expansion of grocery stores is informed by insight provided by teams of location planners and the use of location-based analytics, and this has become widely used by retailers (Birkin et al., 2002, Birkin et al., 2017, Reynolds and Wood, 2010). The use of location planning techniques to aid strategic decision-making became increasingly common and it is widely used across multiple commercial sectors today (Cheng et al., 2007, Hernandez, 2007, Wood and Reynolds, 2011a). Location planning is a broad term used by retail industries for techniques used to model consumer behaviour and to estimate store revenue. It is used to address issues such as store performance, store location planning, the impact of competitors or changes in consumer behaviour. Subsequently, the outputs are used to support decision-making, which can lead to increased turnover and better profit margins by taking the most financially viable options. There are a range of location planning techniques available to support location planning teams, which vary from intuition and experience, to site visits or more objective and practical methods arising from statistical models for instance, such as spatial interaction models (SIMs) or regression analysis (Reynolds and Wood, 2010, Wood and Reynolds, 2011a).

One of the foremost techniques utilised by the grocery sector from the portfolio of tools is the SIM. This is discussed in detail in chapter three. It can be used to estimate store revenue, monitor store performance and provide insight into the interaction between supply and demand. The use of SIMs has become increasingly practical through the increased availability of data (Birkin et al., 2010, Wood and Reynolds, 2011b). The SIM represents the location-planning tool that this research focuses on; a justification of use of the SIM for location based modelling in the grocery sector is provided in chapter 3. The increased competition for retailers to locate attractive and profitable sites has heightened the importance of location planning and accurate store revenue predictions (Wood and Reynolds, 2013). For instance, in the wake of the revised Planning Policy Guidance (PPG6) in 1996, gaining planning permission to develop new stores became increasingly difficult and led to a shift in strategy by grocery retailers; they began to seek out new markets and investments and were increasingly interested in the relationship between supply and demand, and contributed to the rapid growth of the convenience market (Birkin et al., 2002, Hood et al., 2015, Wood et al., 2006). The convenience market expanded rapidly between 2003 and 2012, resulting in the shift from more traditional ‘out-of-town’ supermarket sites to smaller format stores located in more diverse locations, such as city centres, suburbs or stations (Hood et al., 2015). Convenience stores typically cater to 'top-up or secondary mission' shops (Wood and McCarthy, 2014) and tend to have different catchment areas to traditional supermarkets and sales are largely driven by localised footfall (Wood and Browne, 2007), which can be highly temporally dependant depending on the location of the store e.g. at a train station. Subsequently, there appears to be limited application of traditional SIMs for predicting convenience stores sales, because traditional SIMs (which only consider residential populations and are typically focused on larger store formats) have limited predictive accuracy. It became evident throughout the research that there was currently a lack of literature that considers spatiotemporal components of demand within SIMs. The limited literature suggests that if time is underrepresented, retail models that fail to capture actual spatiotemporal components (for instance when using a residential only population) could result in a limited representation of observed patterns of consumer behaviour. Failure to recreate the more complex patterns of consumer demand and grocery expenditure, could lead to inaccurate store revenue estimation (Berry et al., 2016, Malleson and Birkin, 2014, Martin et al., 2015, Newing et al., 2013a). There is a clear need for further research that focuses on the spatiotemporal components of supply and demand in the grocery sector and a need to build these into predictive location planning models (Berry et al., 2016, Hernandez, 2007, Newing et al., 2013a). In particular, Hernandez, (2007), and Newing et al., (2013; 2014) indicate that spatiotemporal developments are necessary, not only to help to overcome the current lack of research in this field, but also to demonstrate the positive advances that spatiotemporal developments (in relation to grocery store location planning) can have for retailers and for modelling techniques, when spatiotemporal components are accounted for, and may overcome the limitations and issues associated with conventional SIMs when used for convenience store location planning. Thus, this research (which seeks to better understand the spatiotemporal dynamics of both supply and demand, to ultimately improve location planning models) offers timely and necessary analyses that help to address current gaps in the academic literature. Furthermore, this research will have positive implications for the commercial sector regarding operational decision-making and modelling capabilities (which are addressed in more detail below and in the conclusion).

Spatiotemporal components are important to retail sectors and in location planning modelling because they have a direct impact upon a store. Space and time can affect the total revenue a store generates (at what time revenue is generated and where the sales originate from) and in regard to the different types of consumer shopping. Time impacts consumer behaviour, altering and controlling the patterns and needs of consumers at certain times of day (Solomon, 2013, Solomon et al., 2013). Temporal constraints in demand, in turn, drive fluctuations in stores as they respond to the spatiotemporal components of consumers. Resultantly, space and time have operational impacts on store operations and design; retailers commonly assign store layouts or adjust the diurnal trade operation of a store in response to spatiotemporal components; for instance, offering ‘food-to-go’ type offers and products (for immediate consumption) and in prime shelf locations in stores, which are driven by lunchtime trade (Berry et al., 2016, Rothwell, 2017). Subsequently, it is easy to see the importance of understanding and being able to build spatiotemporal components of both supply and demand into predictive modelling, which this research seeks to address. The potential benefits resulting from improving our understanding of spatiotemporal components and building supply and demand-side fluctuations into predictive models (for retailers and academia) include:

- A deeper understanding of the spatiotemporal fluctuations of store level sales and the drivers behind them.
- A deeper understanding of the spatiotemporal fluctuation of demand in core catchment areas and the impact on store sales.
- Improved predictive models, which account for diurnal spatiotemporal population movements and the shopping habits of different demand types, resulting in more accurate revenue estimations and a better understanding of catchment area dynamics.
- Following the spatiotemporal development of location planning models, the ability to estimate store revenue in response to different temporal scenarios.
- Insight into store performance over time, which could lead to operational decisions regarding store operations that could allow retailers to optimise store formats, layout and products in response to temporal components, such as demand.

This research seeks to address the currently under-researched areas that are widely acknowledged as necessary and important developments, to the academic and commercial sectors, by analysing supply and demand-side spatiotemporal fluctuations in store turnovers; to better understand the drivers of demand and how and why they vary and what the impacts are at a store level. Subsequently, this research seeks to use this insight to inform predictive and location based decision making by building a series of spatiotemporal demand layers into SIMs and to demonstrate the improved detail and predictive capabilities of such modelling. To achieve this, this research makes use of detailed commercial data (discussed in chapters four and five), rarely accessible for academic purposes, through a close partnership with a major UK grocery retail partner. Using their datasets it presents a novel opportunity to analyse spatial and temporal store level data, as well as loyalty card data, presenting detailed insight into actual consumer level grocery shopping behaviour and spatial patterns. This level of access is rare and offers distinct advantages and opportunities for detailed insight over previous location planning studies, which have not had access to commercial data. This helps to ensure that research analyses resulting from this thesis are novel, timely, accurately informed and applicable to the commercial sector.

The study area for this research is West Yorkshire, which is located in the North of England falling within Yorkshire and Humberside and is the largest Government office region in the UK. The total residential population of West Yorkshire is reported to be around 2.2 million (ONS, 2011b) with an estimated total weekly household expenditure on groceries of £61.4 million (ONS, 2015a). It represents a highly competitive market with intense competition and market shares being fought over by the key players in the UK grocery market. The West Yorkshire region represents a broad range of consumer demographics for grocery shopping. Over the past few decades there has been continued growth by retailers in the area (Hood et al., 2015, Thompson et al., 2012) and it represents a dynamic, yet representative region of the UK.

1.2 Aims and objectives

This research seeks to develop our understanding of temporal fluctuations in store sales and consumer demand to address the need for an increased representation of temporal components within location planning research, and subsequently, to demonstrate the increased accuracy of spatial modelling techniques using spatiotemporal data. The overall aims of this research are:

- To investigate spatiotemporal fluctuations of store sales in the UK grocery sector at a store-level and to explore the demand-side drivers of trade.
- To investigate the spatiotemporal geography of consumer demand and develop a series of daytime population demand layers.
- To build and calibrate a spatial interaction model (SIM) to incorporate spatiotemporal components of demand and demonstrate the improved modelling capabilities for location planning and store revenue estimation techniques resulting from the use of spatiotemporal data.

In order to meet these aims, the following objectives are addressed throughout this thesis:

- To review the current state of literature on consumer behaviour in the UK grocery market and recent changes in grocery shopping attitudes (chapter two).
- To review the role of location modelling in the grocery sector, with a particular focus on SIMs (chapter three).

- To explore the temporal nature of trade at a store-by-store level to ascertain the level of sales fluctuation using EPOS transaction data (chapter four).
- To review the patterns of temporal trade occurring in stores and demonstrate the existence of regular and identifiable patterns of trade leading to the development of a novel temporal segmentation of stores via clustering (chapter four).
- To use the temporal sales profiles of clusters to relate regular and identifiable fluctuations in sales with specific spatiotemporal demand-side drivers of trade (chapter four and five).
- To explore the spatiotemporal geography of grocery demand using loyalty card data (chapter five).
- To produce and calibrate a SIM used in conjunction with spatiotemporal demand layers that can be used to estimate grocery store revenue and account for daytime population movements of transient consumer groups (chapter six, seven and eight).
- To generate a series of spatiotemporal demand layers at LSOA level incorporating daytime spending and behaviour (chapter seven and eight).
- To demonstrate the improved capability of a SIM that accounts for daytime population distributions, as apposed to residential only (night time) populations, and the 'What-if' capabilities of the model; applying the improved SIM (with its increased representation of spatiotemporal demand) in real world type location planning scenarios (chapter eight and nine).
- To explore the practical application of the temporal segmentation of stores using the cluster sales profiles as a location planning tool to estimate the type of diurnal sales profile a new store will experience (chapter nine).
- To review the impact of the spatiotemporal analysis and the practical implications for location planners and grocery retailers, with regards to the potential benefits and strategic decision making, when applying spatiotemporal components in location-based decision-making (chapter ten).

1.3 Thesis structure

Chapter two is the first detailed review of the literature and it introduces the UK grocery market, covering the recent history of the sector and its development up to present day in relation to consumer behaviour, how this has changed and what the driving factors of change have been. Ultimately, the chapter examines the relationship between supply-side and demand-side changes and the ongoing impact that this has had on the grocery market.

Chapter three presents a detailed review of the location planning literature, paying specific attention to the commercial application of location planning modelling, in regard to the reasoning and types of approaches utilised in the UK grocery sector. Specific attention is focused on the application and theory of the spatial interaction model (SIM), providing evidence (from a commercial perspective) for the justification of the decision to focus on this specific tool. The chapter introduces related spatiotemporal modelling approaches and identifies the current issues surrounding these approaches in store location planning. The chapter identifies the commercial need for a better (and current lack of) representation of spatiotemporal components in established location planning methods; providing the justification for this research's applied nature, and highlighting the importance of an applied approach so that it is applicable within grocery location planning teams.

Chapter four addresses the first of the main thesis aims, which is to investigate the spatiotemporal fluctuations of store sales in the UK grocery market and the demand-side drivers of trade at a store level. Utilising a detailed commercial dataset that provides temporal sales records (given by the commercial partners) the chapter provides a detailed analysis of the temporal data. The identification of a series of distinct sales profiles allows the analysis to group stores according to the time of day sales are made. This novel development identifies that traditional segmentation techniques for stores appear to offer limited temporal insight. Moreover, the analysis helps to identify specific temporal patterns that are related to specific demand types (thus addressing the aim), improving our understanding of store level sales and the temporal drivers of the sales patterns that occur in the grocery sector.

Chapter five builds on the evidence that specific demand types drive distinct sales profiles in store revenue at certain times of day, which were noted in chapter four. Thus, the chapter seeks to address the second of the main thesis aims; to examine the spatiotemporal geography of consumers and generate a series of spatiotemporally defined demand layers to apply in a SIM, which to date has been relatively underrepresented in location planning literature. Using a loyalty card dataset, the chapter first correlates the theories regarding temporal demand types with spatiotemporal characteristics identified in chapter four. The chapter develops our understanding of spatiotemporal consumer geographies, which is then used to inform and generate a series of spatiotemporal demand layers that correspond to the observed consumer behaviour patterns. This analysis is used in chapters seven and eight to improve the representation of temporal components of consumer behaviour within SIM based location planning decision-making.

Chapter six focuses on building a detailed and disaggregated SIM for West Yorkshire, calibrated using observed commercial data. This chapter focuses on the traditional SIM and then examines its performance across all store types.

Chapter seven (and eight) address the final thesis aim, to build and calibrate a SIM, which incorporates spatiotemporal components and to demonstrate the positive impact on SIM based location-planning decision-making. Chapter seven focuses on the incorporation and effect of a spatiotemporal workplace population, representing the transient and daytime distribution of consumers who shop from work. The chapter concludes by demonstrating the effect of accounting for more realistic daytime population distributions on store revenue estimations, and compares this to the residential only SIM model built in chapter six.

Chapter eight builds on the spatiotemporal developments developed in chapter seven, and introduces the additional spatiotemporal demand layers identified in chapter five. The model revenue estimations, which seek to be a better and more realistic representation of typical daytime populations, are compared with the residential only model to demonstrate the improved capabilities for SIM based location-planning decision making. The model is also compared with observed consumer interactions; this is to ensure that temporal components of demand and behaviour, and the related impacts to store level sales fluctuations, are robustly represented in the daytime SIM.

Chapter nine will seek to employ the spatiotemporally informed SIM, developed in chapter eight, in a series of 'What-if' location planning scenarios. The chapter focuses on the temporal capabilities of the SIM, demonstrating the positive effects that spatiotemporal data has on a location planners' ability to make informed decisions regarding store planning and the additional level of detail that is on offer for the SIM based approach. Chapter nine also demonstrates the potential practical application resulting from the temporal analysis in chapter four. Using the catchment area characteristics of a store it is possible to predict a store's typical diurnal sales profile according to which one of the four temporal sales profile clusters (developed in chapter four) a potential new store would belong to, based on the catchment area spatiotemporal characteristics.

Finally, chapter ten summarises the findings of the thesis and draws some overall conclusions regarding the contribution of the work in relation to the commercial sector and the implications for location planners, as well as reflecting on the contribution of this research to the academic literature. This chapter will also assess the findings of the research in relation to the aims outlined above and the extent to which each one has been met. The potential of the research will be considered critically in this chapter. It will also discuss opportunities for future research necessary to overcome any remaining, or newly identified, problems and limitations.

Chapter 2 – Changing Consumer Behaviour in the UK Grocery Sector

2.1 Introduction to changing consumer behaviour

The aim of this chapter is to provide insight into the changing trends of consumer behaviour within the UK grocery sector. This draws on changing market conditions, legislative impacts, consumer demographics and the changing roles of supermarket retailers in the UK over recent years and highlights both the cause and consequence of the changing patterns of consumer behaviour in this influential market. The general concept of consumer behaviour is recognised as an ongoing process, whereby behaviour leading up to and preceding a purchase also qualifies as part of consumer behaviour (Solomon, 2013, Solomon et al., 2013). Avery et al. (2013) defined it as a set of activities undertaken by an individual [a *consumer*] in the process of purchasing and consuming products. Within this conceptual framework, a consumer is any person involved in the acquisition of any type of product or service available to them. The product refers to any form of physical or non-physical product or service available, to or sought after by, the consumer (Avery et al., 2013). In this instance, the notion of a consumer's behaviour, in the broadest sense, can be identified as a series of both physical and mental activities. It is important to reiterate that present day definitions are no longer limited to the specific act of buying and using products, but that the physical and mental activities now include time leading into, and long after, the physical act of purchasing an item. Therefore spatiotemporal components become highly relevant.

The UK grocery industry is a very competitive and important retail sector, influencing public conceptions of place and space in the UK, which has resulted in public policy changes to keep control of the market (Birkin et al., 2002, Hood et al., 2015, Wrigley, 1994). It is clear from the literature that the operation of grocery retailing has undergone significant transformations over the last few decades (Keh and Park, 1997). For the purpose of this research, key periods of transition within the grocery market and identified consumer behaviour traits have been clustered in order to tell a concise story. It should be noted that considerable amounts of temporal overlap between each period occur, and in some cases the analogy of 'the chicken and egg' is appropriate in relation to which is driving changes in the grocery sector i.e. is it demand or supply side dynamics? While it proves useful for gaining insight and accessibility to the literature, by sectioning this review, it is important to be mindful of the interconnected nature of consumer behaviour and the drivers influencing change. The intended goal of this chapter is to provide background on the UK grocery market and various types of consumer behaviour (and the determinants of these behaviours). Firstly, the review will cover supply-side developments over recent years and how these have impacted on consumer behaviour in the grocery market. Following this, attention will focus on behavioural changes caused by a range of socio-economic and demographic variables, how they have developed recently and the subsequent impact on store performance. To reiterate the previous point, both supply and demand-side factors clearly go hand in hand in both the development of the supply side as well as influencing consumer behaviour and are of equal interest.

2.2 Retailer growth

The aim of this section is to first situate and discuss the supply side changes that have occurred throughout the substantial market evolution of the last few decades, discussing the subsequent changes in consumer behaviour and patterns of consumption during this period of transition. As the grocery market moved into the 1980s, combined with rising standards of living and improving affluence, patterns of consumption began to evolve, with market dominance being gained through development of product ranges. As consumer behaviour, lifestyles and tastes began to change (such as an increased focus on healthy eating), providing far greater choices and quality of products in supermarkets became commonplace in response to this demand (Burt and Sparks, 1994, Fernie et al., 2010, Harris and Ogbonna, 2001). During this period of transition, the grocery sector underwent changes in operation from the traditional small independent store formats into mass-market superstores focused on scale, efficiency and range. This transformation is highlighted, for example, by the number of 'smaller' CO-OP outlets which dropped from its peak in the mid-1950s of over 30,000 stores to 4730 in the late 1990s (Hallsworth and Bell, 2003). Taking advantage of changing patterns of spending and shifts in consumer behaviour is what ultimately helped lead to a rise in, and domination of, a few market leaders which, at the time, experienced fast growth in response to these changes due to rapid expansion plans (Burt and Sparks, 1994, Burt and Sparks, 2003b, Harris and Ogbonna, 2001).

2.2.1 The Golden age

A noteworthy period of transition in the grocery sector was labelled the 'golden age' (Wrigley, 1987), spanning the late 1970s and early 1990s, it involved the aggressive development of UK store networks (Burt and Sparks, 2003b, Hood et al., 2015, Wrigley, 1994). Due to the rapid growth and subsequent financial success, retailers were operating in an increasingly oligopolistic market and the decreasing power of the manufacturers meant the large retailers developed massive gains in profit, and market and political power (Wrigley, 1994, Fernie et al., 2010, Harris and Ogbonna, 2001). Increasing profits, vast economies of scale and relentless investment and expansion programs, were the result of large out-of-town developments. Retailers were able to operate larger stores at lower costs and the top five food retailers held 61% of the market share at the start of the 1990s, further accelerating growth (Marsden and Wrigley, 1996). This developed an intensively competitive expansion programme with multiple retailers fighting for new developments and paying millions per site in order to stay ahead of the competition (Wrigley, 1994). Burt and Sparks (2003b) noted that this rapid expansion for retailers, created the phenomena described as the ‘spiral of growth’ in which increased sales led to increased sales density and lower operating costs, returning higher profit margins. In turn, this opened up further investment opportunities, repeating the cycle at an increased scale. Those unable to follow this business strategy were reduced to a lower position in the market (Burt and Sparks, 2003b). By the end of this period, the top food retailers were investing billions per annum in expansion programs with consistent rises in profit and net margins reaching between 7% and 9% (Burt and Sparks, 2003b).

However, the period of growth and massive profit margins began to slow down, with early warning signs emerging around 1994 (Burt and Sparks, 2003b). Wrigley (1994) describes how over-commitment to corporate growth and over-paying for sites resulted in huge financial depreciations: for example, profits at Tesco in the space of one year dropped 22% and Sainsbury’s undertook write-offs of over-valued property and sunk costs to the value of £365 million. This downward shift in profits and market shares could be attributed to the impacts caused by three significant factors: (i) the spread and impact of the ‘deep’ discounters’ entry into the UK market; (ii) overestimates in property valuations and the effects of sunk and exit costs; and (iii) public unease about the dominance and power of the grocery retailers, and the eventual introduction of PPG6 in 1993, later revised in 1996 (Wrigley, 1994). A potential fourth cause was the feeling that saturation of the UK market was imminent.

Firstly, the strategies adopted by the top retailers during the golden age in the UK left gaps for market entry at the low end range of the sector. The result was that the UK became an attractive and profitable opportunity for discounters such as, Aldi, Lidl and Netto who began to expand their operations into the UK. Initially, the top retailers ignored the entry of the discounters as it was deemed to be a separately operating market. Discounters began to grow, increasing competition in more vulnerable ‘second tier’ retailers in the UK already struggling to remain competitive, such as Kwiksave and Asda. Burdened with high debt in the 1990s, Asda was threatened at the lower end of the market by the expanding discount retailers (Hood et al., 2015, Wrigley, 1994). In response, retailers (Asda being a prime example of this) refocused their business strategy, concentrating on price competitiveness, and in some cases converted stores and fascias into discount superstores, undercutting the top retailers competitively. The consequence of impact caused by the expansion of the discounter market was finally acknowledged by the top retailers of the time, expressing how profit margins were indeed under threat from the severe price competition of discounters, and they announced the need to respond with long-term price adjustments in order to minimize the loss of profits and maintain customer loyalty (Wrigley, 1994).

Secondly, during the golden age, prices paid by retailers to purchase sites were driven up in an attempt to outbid competitors for attractive sites, with average prices rising by around 30% (Wrigley, 1996). The valuing of these property assets at cost (which was common place at this time), although initially sustainable became less feasible with massive amounts of deflation in property prices starting in the 1990s (Birkin et al., 2017, Wrigley, 1994). Large proportions of the grocery retailers' property portfolios were affected, resulting in high depreciation levels (Wrigley, 1996). Retailers, began to depreciate the value of their assets and slowed store expansion programs, acknowledging that they had paid premium rates for sites during this period (Wrigley, 1994). In order to compensate for the sunk costs retailers applied corrections to pre-tax profits to account for the depreciation of land and store values. Tesco, for example, wiped off a value of £68 million per annum to meet the current property values during the 1990s (Wrigley, 1994, Wrigley, 1996).

Thirdly, was the impact of public concern about growing retailer power and dominance. Following legislation such as the 'Food Safety Act 1990', the power and responsibility for quality control were delegated to the food retailers (Marsden and Wrigley, 1996), accentuating the state's dependency on the need for continued economic growth of retailers necessary to support such changes in legislation (Wrigley, 1994). However, Wrigley (1994) describes how this relationship, which was perceived to privilege already dominant retail leaders, led to increasing unease about the monopoly power existing in the grocery sector. Consequently, increasing criticism of the current market conditions and dominance of these few corporations, and the lack of real competition in the grocery sector, added to this unease (Marsden and Wrigley, 1996). The emergence of the discounters, rather than alleviating this growing unease of power, highlighted the control and privileged relationships held by some of the top retailers. For example, through a number of complaint cases to the Office of Fair Trading (OFT), depicting the struggles of new entrants to the market and reporting suppliers' refusals to trade with them due to existing relationships with top retailers, the situation was intensified (Marsden and Wrigley, 1996, Wrigley, 1994). Retailers were under pressure to reconcile their strategies and engage with the market outside of their own private interests and business operations (Marsden and Wrigley, 1996).

In partial response to the growing concern the UK Government introduced the revised Planning Policy Guidance (PPG6) in 1996. Gone now were the days of laissez-faire attitudes in planning (Birkin et al., 2017, Wood et al., 2006). PPG6 was designed to limit out-of-town development and promote development in town centres, leading to the enhancement of the vitality and viability of town centres, ensuring a vibrant and competitive market (Department of the Environment, 1996). Consequently, gaining planning permission became a much more regulated process, although the theory that the market had become saturated was in fact highly contested (Birkin et al., 2017), planning guidelines made site acquisition much more difficult. Although some retailers had purchased land (typically out-of-town sites) prior to legislation changes with prior planning permission for those sites, waiting to develop stores at a later date (known as land banking), retail location planning teams were forced to address store development in different ways, such as different formats for expansion and a move into new geographical locations (Birkin et al., 2017). Through improvements and changes to location model designs and in response to the revised PPG6, store location teams were able to suggest new methods of expansion, identifying opportunities within the tighter legislation using analysis methods, such as stores per head or floor space per head, identifying areas that were still suffering from under-provision of grocery services (Birkin et al., 2017).

One of the innovative ways that supermarkets responded to greater planning restrictions was to enlarge previously opened stores, offering larger product ranges, particularly in non-food products. While increasing the product range helped to increase store turnover and improve attractiveness, research suggests that preference for food products was still of central importance for consumers (Wood et al., 2006). Therefore, retailers avoided reducing floorspace for food products, and instead created new floorspace for the extended product lines. However, due to legislation preventing exterior changes to existing stores (section 55(2)(a) of the Town and Country Planning Act 1990), new and innovative methods were adopted. For example, Asda identified a loophole in the legislation which allowed development and extension projects within the store without being affected by restricted planning permission. Asda began to experiment with floorspace expansion through the building of mezzanine floors (Wood et al., 2006). Large supermarket sites were thus able to expand their floorspace considerably (Asda's Tamworth store was able to add an extra 30,000 sqft to the already 55,000 sqft store). Unsurprisingly, this loophole caused renewed concern about the effects on and competition to high streets and town centres, so changes were again proposed to the Town and Country Planning Act, to limit increasing internal floorspace to a threshold of 200m^2 (Wood et al., 2006).

During this time period, Tesco increasingly became interested in industrial 'brownfield' sites and deprived housing estates, which were previously avoided due to the perceived limited disposable incomes of the surrounding catchments (Wood et al., 2006). Through its so-called regeneration schemes, Tesco were able to drive store expansion programs forward, gaining planning permission in return for local regeneration projects. In partnership with local authorities Tesco emphasised the potential benefits to the population and sites, (although retailers openly described their motives in regeneration schemes as self-enlightening), bypassing tightened planning regulations (Wood et al., 2006). Benefits included the contribution to employment opportunities to the local workforce and subsequent skills and training provided to the local population, adding opportunities of long-term employment as well as providing essential services and quality food shopping opportunities (Wood et al., 2006, Birkin et al., 2017). Gaining planning permission through regeneration schemes has allowed Tesco to gain significant increases in floorspace, adding more than 1 million sqft, resulting in significant increases in market penetration (Birkin et al., 2017, Wood et al., 2006). Other strategies included the adoption of smaller store formats (between 10,000 and 20,000 sqft) which could be opened in smaller towns or suburbs; for instance in the case of Tesco, the development of the 'Compact' store. Capable of operating at lower profit margins but requiring smaller investments, these were suitable for smaller catchments and in line with the tighter planning guidelines (Wrigley, 1994). Furthermore, there was a sizable investment into the convenience store format, which is discussed later.

Finally, following the end of the Golden age, there was a perceived belief that the grocery market in the UK was close to saturation. This was caused by the belief that due to the top retailers dominant market share, there was very little room left for growth of other retailers and that few viable sites for new development existed. Concern over saturation was first noted in the late 1980s and again in the early 1990s (Duke, 1989, Jones, 1982). In addition as several senior executives in the grocery industry (Thompson et al., 2012), announced that saturation was very imminent within the market or had already occurred. The belief in saturation led to grocery retailers slowing rates of expansion as growth seemed no longer so easy (Burt and Sparks, 2003b), instead seeking continued profit growth through cost savings, international growth or refurbishments of the existing network (Birkin et al., 2002, Guy, 1996). While many of the assumptions were made with very little analysis, this speculation still had an effect on business operation. It was not until the mid-to-late 1900s that detailed analysis of provision found that saturation of the grocery industry was unlikely and proposed that it could only ever be a local phenomenon (Guy, 1996, Langston et al., 1997, Langston et al., 1998).

Similarly, consumer behaviour underwent similar changes as a result of these supply-side transformations. With the mass expansion both in the number and size of out-of-town stores, an extensive consumer choice was now on offer. Consumer behaviour was observed to take advantage of the range of products now available at one site. Patronage of shoppers shifted from the traditional town centre or high street to the large out-of-town stores. Basket sizes grew, with consumers spending more and buying in bulk, reducing the frequency of shopping trips with the norm becoming a 'one-stop' weekly shop (Akehurst, 1984, De Kervenoael et al., 2006, Hunter, 2004). In addition, the rise of car ownership, ability to transport more items, as well as extensive parking opportunities encouraged this, with consumers travelling further for better facilities. Intrabrand competition between retailers and manufactures also allowed stores to attract customers from far away with cheaper products, bulk discounts and extensive product ranges (De Kervenoael et al., 2006, Hunter, 2004, Wood et al., 2006). At the time customers favoured the large-scale retail opportunities, seeking services and products that smaller local stores were unable to offer (Clarke, 2000, Burt and Sparks, 2003b). Retailers began extending product lines outside of food, providing clothing ranges, e.g. George by Asda, electronics and entertainment aisles in an attempt to cater for the developing demands of consumers. The resultant effect was that consumers were offered multiple product ranges by supermarkets (further drawing spending away from traditional retail formats) and developing considerable brand and product loyalty between the customer and retailer (Akehurst, 1984, Burt and Sparks, 2003b, Elliott et al., 2012, Penton-Media, 2001).

Following the decline of the major retail 'golden age' period the behaviour of the UK grocery consumer was observed to change once again: as discussed in section 2.4. Following this, in the 1990s, consumers became attracted by value-for-money and basket size began to decline as consumers sought out the best discounts, often splitting shopping trips across multiple stores and making use of the development and expansion of the discount market, further reinforcing the point that changes in consumer behaviour and supply, which for the purpose of this research are broken down by theme, are intrinsically linked to one another.

2.3 Socioeconomic and demographic driven behaviour

There are various socioeconomic and demographic factors which contribute to the differences in the behaviour of individuals and these have received substantial attention across different research fields (Avery et al., 2013, Birkin et al., 2002, Solomon et al., 2013, Thompson, 2013). Nevertheless, while the segmentation of consumers by socio-economic and demographic variables has remained a relatively constant commercial practice, populations and attitudes are continually evolving, resulting in shifting patterns of consumer behaviour as individual aspects of consumers change. For the purpose of this research, common and influential social-economic and demographic variables will be discussed, providing insight into the impact on consumer behaviour and the resultant changing dynamics in the UK consumer population in regard to the grocery sector.

2.3.1 Social class, Income and Expenditure

Although social class should only be used as a proxy for income, as considerable difference particularly in the higher social class categories exists, generally speaking individuals of similar social class are likely to have approximately similar incomes (Solomon, 2013, Solomon et al., 2013). Consequently, they are also likely to demonstrate similar habits, tastes and lifestyles which can influence consumer behaviour and spending. Household and individual income have been shown to highly influence consumer expenditure levels across all services and commodities (ONS, 2015a, Piacentini et al., 2001). Evidence supporting this can be seen through the Office for National Statistics Family Spending: Living cost and Food Survey report which segments spending by output area classifications as shown in figure 2.1 (a). It is possible to breakdown income dependent behaviour by higher and lower income earners and as shown in figure 2.1(a), higher earning social class groups generally spend more money overall per week. This is particularly noted with higher spending in recreation, transport and food services.

Although, when viewed proportionally as a percentage, shown in figure 2.1(b), lower income households spend more of their weekly expenditure on food than their wealthier counterparts.

Typically traits for high-income earners [in the grocery sector] are linked to increased spending levels and a stronger attraction to better quality, healthy and discretionary products. In contrast the typical consumer behaviour of a low income earner is linked to price dependant products, buying cheaper ‘essential’ items often in place of healthier or higher quality products. In these instances households have been noted to adopt sophisticated ‘price-driven’ and ‘economic’ shopping strategies (Jetter and Cassady, 2006, Piacentini et al., 2001, Rohm and Swaminathan, 2004). This makes marketing and segmentation of consumers by income an important variable for retailers to understand, and as income and social class fluctuate, spending habits are likely to change too. At the beginning of the past century the general observed pattern was an increasing level of income and consequently consumers shopping habits have shifted, matching increased disposable incomes (Birkin et al., 2002). Within the UK this trend was seen up to 2008 when income and subsequently disposable income peaked and then began to decline (shown in figure 2.2). This downturn in income was partly linked to the global financial crisis (starting at the end of 2007 and culminating in a full financial meltdown in 2008) with several billion pounds worth of bailouts to the UK banking system occurring throughout the period (BBC, 2015, Havemann, 2015, ONS, 2015a). Figure 2.3 provides a breakdown of changing consumer spending habits in the UK over a ten-year period (ONS, 2015a), identifying the shifting patterns in spending preferences for overall expenditure by commodity or service in UK households. While income and spending in most service areas has declined over this period, *housing, fuel & power, recreation & culture and clothing & footwear* spending actually increased. Expected patterns of spending between income and social class groups appear to remain consistent with higher income earners spending more money than less affluent counterparts. However, the overall decline in spending is likely to be linked to a continued ‘price conscious consumer behaviour’ following the 2008 economic downturn (and shifts in the perception of what is now consider an essential product) not to mention the increased discount market that was now available.

Figure 2.1(a) - Shows the breakdown of weekly expenditure between different OAC supergroups, Source: (ONS, 2015a)

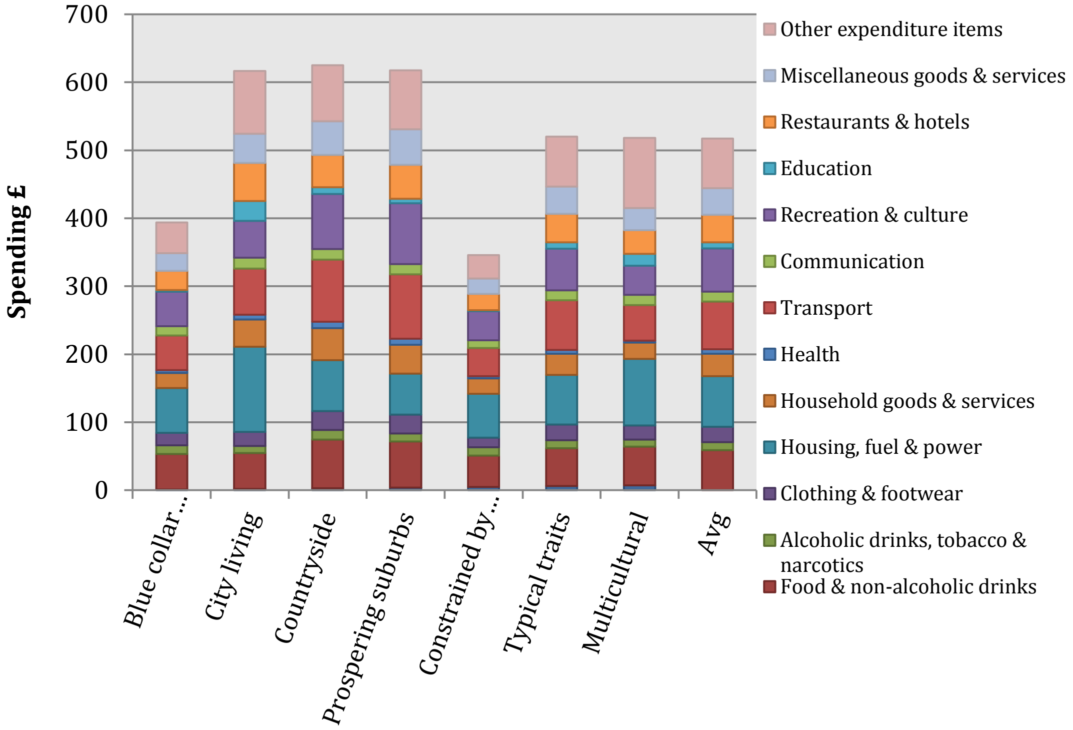


Figure 2.1(b) - Shows the breakdown of weekly expenditure between different OAC supergroups as proportional spend, Source: (ONS, 2015a).

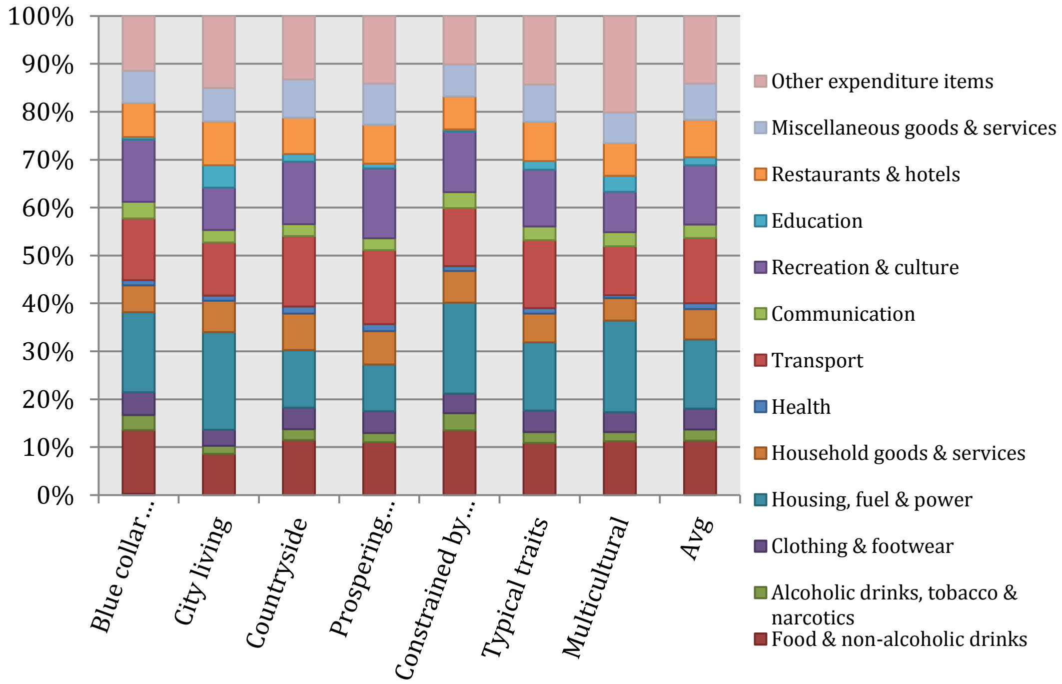
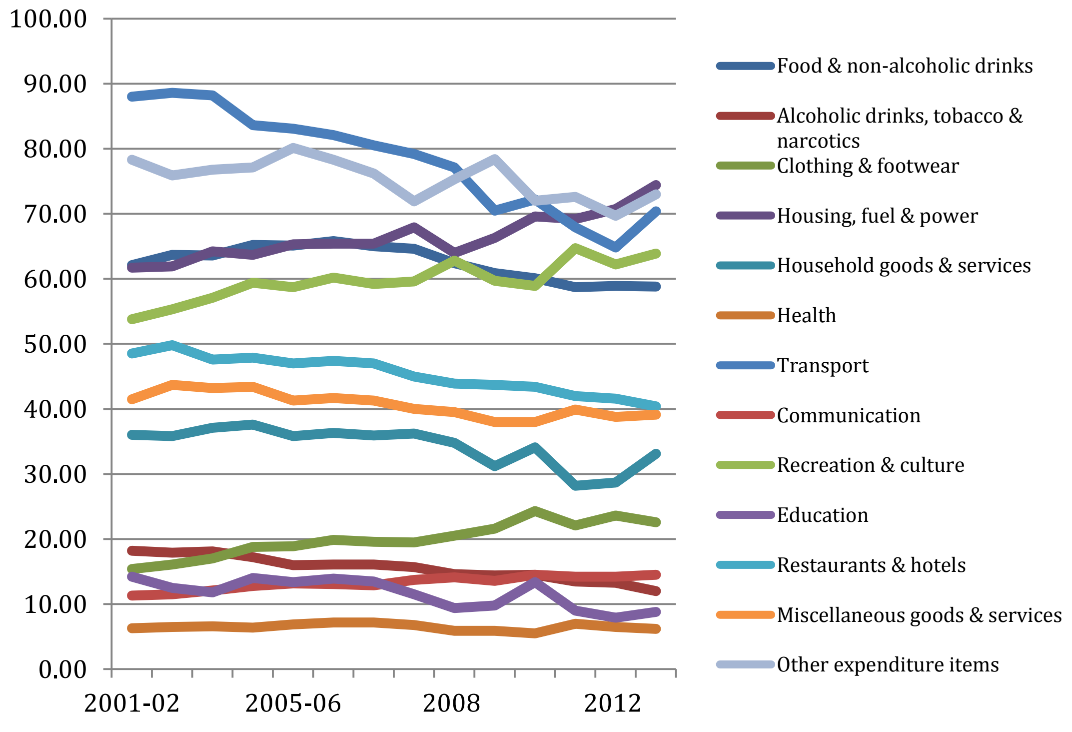


Figure 2.2 - Shows weekly gross and disposable income per average UK household over a ten year period, Source: (ONS, 2015a).

Figure 2.3 - Shows the breakdown of average weekly expenditure in the UK between different commodities and services over a 10 year period, Source: (ONS, 2015a).



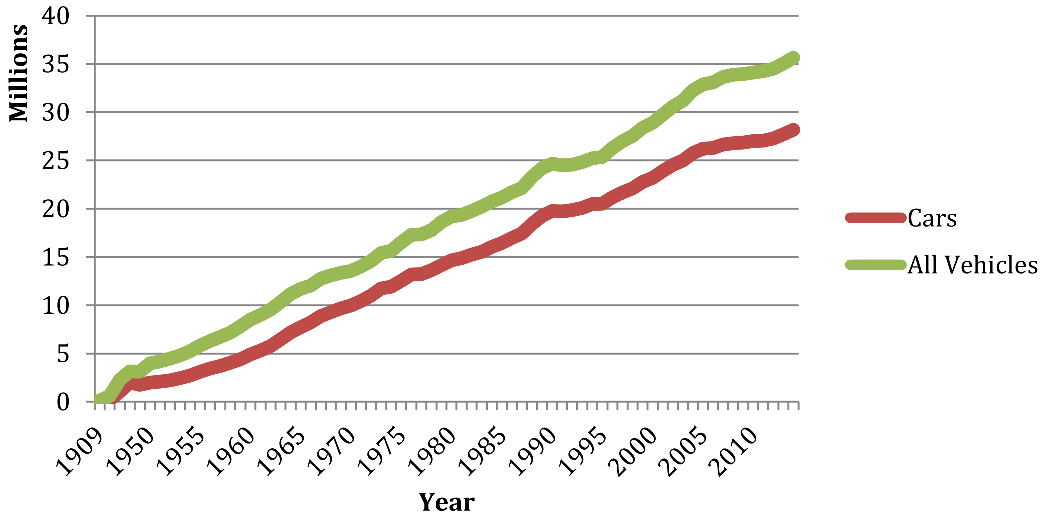
2006 (a) From 2001-02 to this version of 2006, figures shown are based on weighted data using non-response weights based on the 1991 Census and population figures from the 1991 and 2001 Censuses.
2006 (b) From this version of 2006, figures shown are based on weighted data using updated weights, with non-response weights and population figures based on the 2001 Census.

2.3.2 Consumer Mobility

This is commonly associated with spatial effects like the issue of distance and access to transport, levels of car ownership, and has been widely discussed as a major influence on consumer mobility (Birkin et al., 2002, Bowlby, 1979, Thompson, 2013). Consumer mobility is also linked to job/family characteristics, income or physical problems like age and health, highlighting the connected nature of variables that impact upon consumer behaviour (Bowlby, 1979). However, for the purpose of this research attention will be focused on car ownership and the changing levels within the UK and resulting impacts on the British grocery shopper. Car ownership impacts upon a consumer's mobility and their choice and ability to travel further to bigger, better or more suitable supermarkets in addition to being able to carry more items. When catchments are extended through increased mobility, consumers are able to seek out better quality and extended product lines, whereas individuals with reduced and restricted mobility are limited to store offers in the immediate locality. High mobility is often associated with younger, more affluent individuals and who subsequently experience a better quality of life.

Consequently, customers with lower levels of mobility (commonly elderly or lower income households) suffer from poor access to grocery stores, either due to poor local choice or the inability to travel further to better service provisions, which acts as a form of social exclusion (Piacentini et al., 2001). These areas are often referred to as '*food deserts*' i.e. in areas containing fewer choices for grocery shopping and as a result, consumer behaviour in these locations is linked to increased levels of unhealthy diets and a strong relationship with obesity and poor health (De Kervenoael et al., 2006, Hallsworth and Bell, 2003, Wrigley et al., 2002). Research has found that consumers with poor mobility often rely on others for their food and shopping requirements (Piacentini et al., 2001). Figure 2.4 shows levels of (licensed) car ownership for the past century in Great Britain. There is a clear trend demonstrating an increase in the number of licensed cars in the country. Consequently, consumers are able to spend additional time travelling, are capable of buying more products on a single trip and exhibit multipurpose trip behaviour, travelling to multiple retail destinations during single shopping activities with increased accessibility to bigger and better products (Goldman and Hino, 2005, Leszczyc et al., 2004, Rasouli and Timmermans, 2013a, Piacentini et al., 2001).

Figure 2.4 - Shows the number of licensed vehicles in Great Britain between 1909-2014 (Grove, 2015).



2.3.3 Age

The age of a consumer has a considerable impact on their needs and wants. It should be noted that while individuals of the same age obviously differ in many ways, consumers have been shown to exhibit consistent values and common cultural experiences within age groups, identifying shared behaviours and trends across the consumer age spectrum (Avery et al., 2013, Solomon, 2013, Solomon et al., 2013). Therefore, understanding the types of behaviour shared by specific age groups is a crucial component for retailers in marketing to tailor the shopping experience for specific consumer age groups (Solomon, 2013, Solomon et al., 2013). While often considered disadvantaged and socially excluded customers (due to mobility and accessibility complications), the older population represents a substantial consumer group for grocery retailing and they have been found to favour higher-quality products, good service and convenience when undertaken grocery shopping activities (Kohijoki, 2011, Piacentini et al., 2001). The older generation when grocery shopping demonstrates a tendency to favour stores within close proximity to home as well as larger supermarket stores. They are also noted as having a reduced frequency of weekly grocery shops. Older customers are also observed to not consider online shopping as an accessible option with few considering ordering groceries via the internet, favouring a physical brick-and-mortar store experience and demonstrating a high level of brand loyalty (Kohijoki, 2011). While the elderly have been noted to consider themselves as not disadvantaged, evidence suggests that their shopping habits and behaviour are highly susceptible to changes in the grocery market and as a result likely to suffer from the closure of local stores with issues of accessibility becoming increasingly important (Kohijoki, 2011, Piacentini et al., 2001). On the other hand, younger customers tend to be less disadvantaged in terms of access (with higher mobility levels) and so are able to travel further and make more frequent visits to supermarkets (Burt and Sparks, 2003a, Kamarulzaman, 2010, Kenhove and De Wulf, 2000). They have also been shown to have a higher uptake of online shopping, relying on the internet to make purchases and favouring time-saving behaviour (Kamarulzaman, 2010, Kenhove and De Wulf, 2000).

The beliefs, experiences and values of a consumer are considered to be relatively consistent according to age, developing and evolving throughout life (Solomon et al., 2013). Behaviour and values are shared by persons of the same age, becoming important or attractive and then fading again as age changes (Avery et al., 2013). Therefore, while the physical population of an area may stay the same, it can be argued that the behavioural attitudes and habits exhibited over time by an ageing population changes. Therefore in order to understand the changing dynamic of consumer behaviour and the impact on the grocery sector, the best approach is to examine the current, and projected, spatial representation of the UK age profile. Using census data it is possible for retailers to assess how the age profile has changed, with the current age distribution and predictions for the future contributing to the understanding of changing consumer behaviour in the UK population, and the types of services that will be most profitable.

For instance figure 2.5 shows that while a large proportion of the UK is represented by young adults between 20-35, there is an ageing population with 65+ showing a continuous increase since the 1970s. The older population is set to continue to increase within the next 10-20 years. Coupled with improving healthcare systems and living standards resulting in an overall increase in life expectancy (Birkin et al., 2002), it is likely that we will continue to see rising population levels particularly within the over 65 age group, and as age profiles change, retailers will have to adapt to ensure they offer appropriate services and support (Kohijoki, 2011). The standard practice for retailers is to ensure a market and services where the needs of each consumer age group are met, providing accessibility and retail environments suitable for all (Solomon, 2013, Solomon et al., 2013). Assuming a reasonably level of accuracy in this data on future predictions (figure 2.6), the current market will need to focus on providing a continued service to young adults as well as increasing the services and products aimed at the more elderly consumer.

Figure 2.5 - Breakdown of the UK population by age, between 1953-2013 (ONS, 2014d)

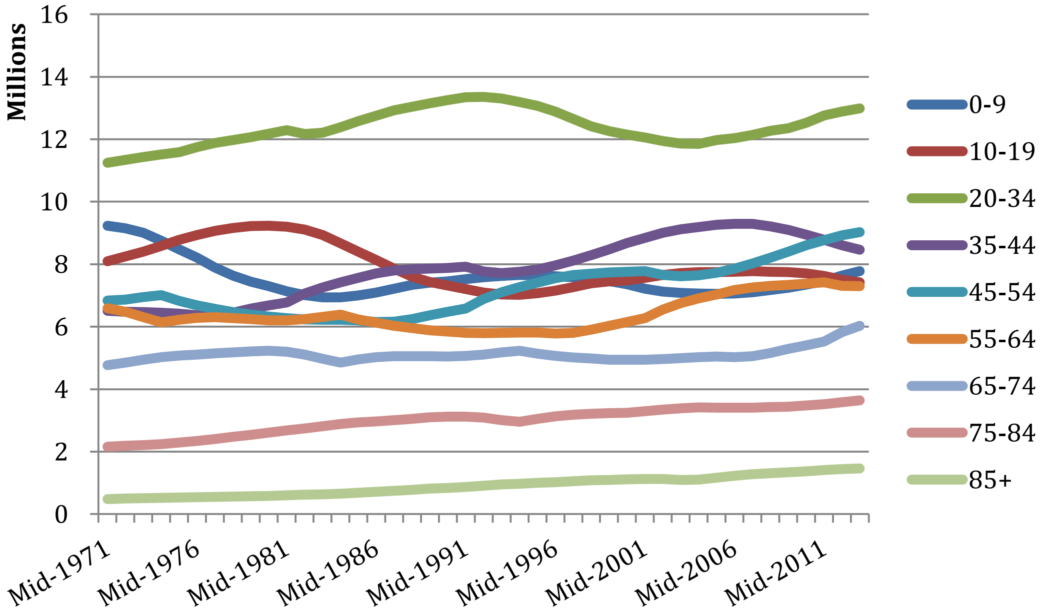
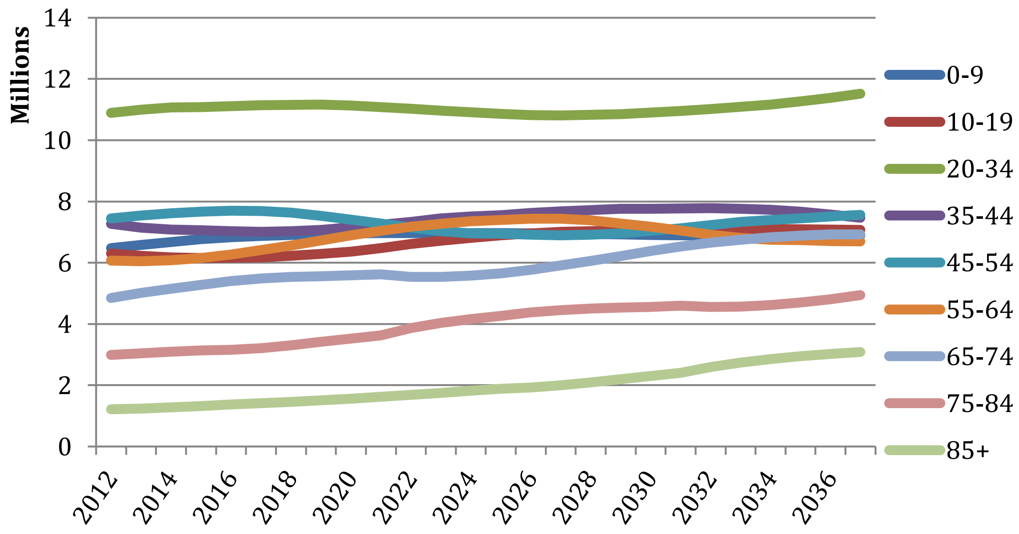


Figure 2.6 shows the predicted population projections for the UK by age, between 2012-2036 (based on 2012 trends) (ONS, 2014c).



2.3.4 Gender

Gender distinctions in retail can be seen from a very early age, with many different products designed and marketed for male and female consumers (Solomon, 2013, Solomon et al., 2013). Traditional female gender roles in the early post-war years suggested that women sought more feminine products, were more concerned with health and how everyone around them felt. Their role was that of the homemaker, with those who did have jobs usually being in subservient roles. Women were typically perceived as being responsible for decision making and shopping when it concerned everyday living and homemaking. Grocery shopping was predominantly seen as the role of the female with many firms marketing products to women and housewives (Avery et al., 2013, Little et al., 2009, Piper and Capella, 1993). Conversely, traditional gender roles for men professed that a man's role was out of the house, working and earning the main income for the household. Men were described as masculine, favouring products that provided intrinsic benefits as well as valuing a product solely based on its utility. The male's role as the consumer was as the 'principle' decision maker and this was reinforced through the belief that men were responsible for purchasing products such as cars and DIY products (Avery et al., 2013, Piper and Capella, 1993). While research undertaken by the UK Food Standard Agency (FSA) in 2007 found that in 77% of households females were still responsible for undertaking the grocery shop (FNS, 2007), modern day social and demographic changes are contributing to new and developing gender roles, resulting in a much more diffuse identity, with changing perceptions and sharing of the traditional gender roles and consequently evolving consumer behaviour (Piper and Capella, 1993, Timmermans et al., 2002). Subsequently, more recent research by the FSA in 2014 found that the responsibility for grocery shopping according to gender had shifted by almost 10%, to 68% for females and 32% for males (Prior et al., 2014). Factors accounting for changing female behaviour are linked to increasing numbers of educated working professional women, and households are now more commonly supported by both male and female income earners, with couples developing equal or reversed gender roles, living in smaller families and having marriages later in life (Avery et al., 2013, Timmermans et al., 2002, Piper and Capella, 1993). Male consumers have been observed to be taking on increased roles relating to the household, adopting more responsibility and consequently grocery shopping is noted to be transforming gender identities with more males voluntarily engaging in grocery shopping with and without their partners (Piper and Capella, 1993, Thompson, 2013).

2.3.5 Ethnicity

Increasing ethnicity levels in the UK, and growing number of diverse cultures, have led to a growth of diverse and deeply rooted ethnic consumption patterns in a multi-cultural population, resulting in newly established consumer attitudes towards food (Hamlett et al., 2008, Solomon et al., 2013, Wang and Lo, 2007). Common ethnic consumption patterns within the UK which effect grocery retailers include the purchasing of traditional foodstuffs of different ethnic groups, such as halal meat or imported spices. As a result, ethnic consumers are observed to have a higher tendency to undertake their 'functional' grocery shopping purchases using traditional independent retailers who specialise in ethnic food products, using major supermarkets as social outings to experience new things and purchase mainstream or bulky household items (Goldman and Hino, 2005, Hamlett et al., 2008). Other observed ethnic consumer behaviour norms for some ethnic groups include the belief that shopping trips made by unaccompanied females must be made locally, limiting accessibility and choice in the first place (Goldman and Hino, 2005). To some extent major supermarket retailers have attempted to evolve and cater for increasing ethnic diversity through different product ranges, stocking increased ranges of world foods and ethnic ingredients (Hamlett et al., 2008). According to the market research company Mintel, the ethnic market represents a substantial proportion of consumer spending valued at £1.4 billion and for example, increased by 6.6% in 2011 alone (Mintel, 2012). Furthermore estimated patterns of immigration to the UK, increasing by approximately 7% between 2005 and 2014 to over 640,000 immigrants per year (ONS, 2015b). An increase in ethnic-cultural minority markets and available products will probably continue to be seen in response to this growing consumer demand, either through further growth and development of specialist ethnic independent stores (Birkin et al., 2017), or through the expansion of major supermarket product ranges.

2.4 Austerity and rise of the discounters

A prominent stage in the price driven transition (of a relatively speaking 'modern' Britain), impacting upon both demand and supply sides, was the period of post-war retailing (Burt and Sparks, 1994). During the war retailers, as well as the grocery sector, were under ration control and consequentially consumer behaviour was also restricted, limiting product choice, weekly spending and frequency of shopping trips. However, in the early years following the end of the Second World War, Britain began to undergo massive changes in social and consumer behaviour with a period of changing retail habits beginning in the 1950s. With the end of the war and austerity and the removal of rationing, lifestyles began to return to 'normal' and the conventional patterns of wartime consumer behaviour were eroded. The amount of products available in the market increased and consumers were free from strict government controls. They had more disposable income than before, and instrumentally, retailers were no longer confined by retail price maintenance (RPM) controls (after 1964), contributing to a booming retail economy with increased consumer mobility and affluence: consumers could now afford more expensive goods and had much more choice (Dashwood, 2013b).

RPM was a common practice whereby manufactures set minimum prices for products that retailers had no choice but to adhere to. In setting a price for products it prevented price wars and market competition through retailers offering discounts and thus maintained an 'equal' market. However, following the removal of RPM in 1964 (Hallsworth and Bell, 2003), increased spending and retail prosperity was particularly seen in the grocery sector with the introduction of the self-service store, developing into what we now consider a conventional supermarket format, which revolutionised grocery retailing in the UK. Self-service stores were first established by Sainsbury's in 1950s London and consumers were now able to shop more freely, enjoying the freedom to shop at their own convenience without the need to queue and speak to a shop attendant, fundamentally altering consumer behaviour of the day. While products were sold at cheaper prices, the increased availability and choice of products ultimately led to consumers increasing their basket sizes and spending more at the benefit of the retailer (Dashwood, 2013b). Again, rising patterns of affluence and mobility in the early 1960s led to further spending becoming more common and patterns of purchasing and shopping behaviour began to transform along with the UK grocery sector. Large scale retail operations were now providing discounted prices and bigger product ranges all under one roof (Burt and Sparks, 1994, Keh and Park, 1997). Major grocery retailers had seen the benefit of self-service formats and by the late 1960s there were over 24,000 stores offering savings through their own bulk buying. Impacts to brand loyalty were prominent, as consumers were now beginning to exhibit a new behavioural pattern, choosing to shop around looking for better products and prices playing off retailer competition instead of remaining loyal to the traditional single store.

In the early 1990s a second major price driven change was caused by the arrival of the discounters in Britain. The development of the discount market was in part initiated due to the operational decisions made by the top retailers during the collapse of the golden age (see section 2.2). The strategies adopted by the top retailers during the golden age development left gaps in the barrier to market at the lower end price range of the market. Much of the golden age expansion was directed to out-of-town sites and away from less affluent suburbs, meaning that discounters were able to exploit this niche market with little to no competition other than the incumbent retailer Kwiksave. The result was that the UK became an attractive and profitable opportunity for large European discounters, such as Aldi and Netto, who began to expand their operations into the UK (Burt and Sparks, 1994, Wrigley, 1994). Initially, the top retailers ignored the entry of the discounters as it was deemed to be a separate market from that in which they operated and thus would not affect their overall gross profit margins. However, following the recession in the 1990s giving discounters their first opportunity of growth, a second period of recession between 2007 and 2010 (Thompson et al., 2012), further strengthened the position of the discounters in the grocery market.

The prediction was that by the mid-late 2000s there would be strong levels of growth in the UK discount sector and by 2008 Aldi, Lidl and Netto (the main discounter retailers) exhibited a collective market share of 6.1%, which was the highest level seen in the UK for the discount market (Birkin et al., 2017). The growth of market shares can be ascribed to rapid expansion programs extending market penetration throughout the UK. Take, for example, Aldi: in 2009 it opened 50 new stores and planned to open a further 29 stores the following year with a total of 510 stores by 2014; similarly, Lidl announced plans in 2013 to double their store portfolio from around 600 to 1200 stores (Birkin et al., 2017), highlighting the increasing presence and success of discounters. Austerity measures in the UK have also partly led to the rise of a different sub-set of discounters offering a mixture of dry/packaged food and general goods from high street locations, such as Poundland, Poundworld and Wilkinson's.

The traditional classification of the discount shopper was associated with the lower income consumer, not restricted by age, but by those consumers who sought and needed to find the most local retailers and at lowest prices (Birkin et al., 2002, Piacentini et al., 2001). However, recent increases in market shares were also linked to the patronage of more affluent shoppers shifting to some extent from major supermarkets to discounters, with an increase in consumers now focused on looking for deals and better value-for-money products (Birkin et al., 2017, Thompson et al., 2012). Research, in 2012 identified that 21% of the discounter Aldi's customer profile was made up from the 'prospering suburbs' OAC (Thompson et al., 2012), emphasizing more price driven behaviour by middle class consumers. Consumers were seeking out the cheaper prices offered by discounters, with value-for-money becoming (once again) a substantial factor in grocery shopping influencing consumer behaviour, with consumers increasingly referred to as *economic* or *smart* shoppers (Groeppel-Klein et al., 1999, Piacentini et al., 2001, Rohm and Swaminathan, 2004, Westbrook and Black, 1985). Today shoppers are more likely to travel to sites with multiple retailers, spending more time looking at products often using multiple retailers to purchase products, ensuring the best deal. Consumers are now exhibiting greater brand-switching behaviour, undertaking shopping and purchasing different products at a range of stores (driven by the price benefits gained and value-for-money achieved by switching retailers and shopping around). The threat created by shifting patterns of spending has been taken seriously by the major retailers, aggressively fighting back lost market shares through the introduction of new low-cost ranges and discounts on day-to-day essentials (Thompson et al., 2012). This has likely drawn some consumer spending back from the increasing discount markets, and consequently has allowed consumers to further benefit, playing off the increased price competition generated between discounters and major retailers.

2.5 The drive for convenience

Following the problems of the 1990s grocery retailers shifted expansion programmes from the large super and hypermarket store formats of the golden age, instead seeking out profits from smaller sites in inner city and suburban locations, thus reverting back to the convenience style format of the pre-war period with small stores on every high street (Dashwood, 2013b). Convenience stores, generally defined as smaller than 3000 sqft (a threshold which allows extended trading hours) (Birkin et al., 2017). The major grocery retailers aggressively entered the convenience market as a result of these policy and planning changes. Coinciding with the tougher controls on the retail market and planning legislation (PPG6), along with significant changes in the lifestyle characteristics of consumers, there was a growing need for retailers to meet the rise in demand for convenience (Birkin et al., 2017, Hood et al., 2015, Wood et al., 2006). Originally the convenience market was dominated by independents, particularly during the golden age of grocery growth (with major retailers focusing on larger superstores). They provided the advantage of location and longer opening hours. The aim of this section is to introduce the convenience market, identify to some extent the supply-side drivers to retailer operations, but also the changes in consumer behaviour that similarly initiated and resulted from the growing demand in the convenience market.

2.5.1 The convenience market

Since the entry of the major retailers into the convenience market there is no doubt that independents have suffered in more recent times. This was particularly seen, for example, with the decline of independent fruit and veg shops, bakers and fishmongers (Birkin et al., 2017). The decline of independents has also been linked to operational problems, such as a lack of business experience, lack of capital and low operating margins along with the complexities of the supply chain adding to the decline in market share of the independents in the convenience sector (Birkin et al., 2017). However, in the current market conditions, not all of the decline can be attributed to store closure, with some independents (due to the increasing pressures and intensifying competition) opting to join symbol groups such as Spar, resulting in a sizable convenience market share of 40% for this category (Birkin et al., 2017). Symbol groups are umbrella organisations under which stores operate, taking advantage of greater economies of scale and buying power, with access to own brand products, logistics and business support from head offices. They provided independents with security and support from the competition of major retailers entering the market. Convenience store numbers operating under symbol groups rose from 6900 in 2000 to 15,491 in 2017 (Birkin et al., 2017, Roberts, 2017), highlighting the pressures on independents attempting to remain open in an increasingly monopolized and competitive market. The current national market share for symbols and independents across all grocery retailers according to the consumer research group Kantar Worldpanel mid-2017 was 2.1% (Kantar-Worldpanel, 2017).

As major retailers were forced by the revised PPG6 to explore new locations and markets in order to continue corporate growth, this growth was also fuelled by the decision of the Competition Commission that traditional supermarket grocery retailing and convenience grocery retailing were two separate markets. Major retailers in the grocery sector have taken this opportunity to rapidly expand into this market, using organic growth alongside acquisitions to rapidly expand operations in relatively short periods of time (Birkin et al., 2017, Wood et al., 2006). The convenience market grew from £19billion in 2000 to £37.7 billion in 2015 and is set to be worth £49 billion in 2019 (Birkin et al., 2017, IGD, 2014). Initially the first major retailer to adopt and expand into the convenience format was the Co-Op, seeking shelter from the fierce competition of the golden age superstore format, and seeking strategic 'first mover' rights (Hallsworth and Bell, 2003). The retailer initiated a tactical review highlighting a re-emphasis on local and convenience stores, and through the acquisition of 600 Alldays stores in 2002 and 121 Balfour stores in 2003 it became a primary retailer in this market (Birkin et al., 2017, Wood et al., 2006). However, the expansion of the remaining major retailers into the convenience market ultimately shifted the balance of power from independents to major retailers and this was initiated by Tesco in 2003 (Wood et al., 2006). Following in the footsteps of Co-Op, Tesco purchased 862 stores through an acquisition of T&S in 2003 operating under the 'One Stop' banner and in combination with rolling out their own brand of convenience stores under the Tesco 'Express' format. Through these actions they significantly altered the competitive landscape of the convenience market (Wood et al., 2006). Others soon followed, such as Sainsbury's and more recently Waitrose in 2011, respectively through their own convenience formats: Sainsbury's 'local' and 'little' Waitrose (Birkin et al., 2017). Through the rapid roll out of major retailers own brand convenience stores (due to attractive profit margins and relative ease in obtaining sites) and continued acquisitions of existing businesses, a substantial and quickly evolving landscape emerged between retailers competing for large market shares in the convenience sector, taking advantage of expansion opportunities and employing high economies of scale for profit extraction (Birkin et al., 2017).

Birkin et al. (2015) describe how traditional location analysis techniques seem to have become redundant for the convenience store market, with the need to adopt simpler quantitative techniques and simpler forms of GIS in combination with more site visits to establish new convenience store locations, although little literature has been published on this subject, often infilling around the catchments of existing superstores. Convenience stores are currently being located among the following locations: high streets, train and petrol stations, rural towns and villages, residential populations such as student suburbs and areas with highly temporal populations such as business districts (Birkin et al., 2017, Hood et al., 2015); providing ease of access, longer hours and convenient shopping for fresh quality products to customers in locations where services might not have previously been provided. Due to the attractive nature of the grocery convenience market, intense consumer demand, current lack of tight planning regulations in this sector and relatively low financial risk, it is likely that competition in this market will intensify with increased pressure applied on store location decisions to find new sites (Birkin et al., 2017). It is reported that the convenience market generated £37.7 billion in the 12 months leading up to April 2015, with a year-on-year rise of 5.1% (IGD, 2015).

2.5.2 Demand and consumer behaviour drivers

It can be argued that supply-side changes have fed demand in the convenience market, supplying consumers with these services and provisions, resulting in changing consumer behaviour and consumption patterns. Equally, on the other hand, it can be argued that changes in the demand for convenience are associated with demographic and lifestyle shifts, altering consumption patterns and consumer behaviour and necessitating the need for the development of the convenience market (Hallsworth et al., 2010, Hood et al., 2015). This section aims to address these demand-side changes, providing insight into shifting patterns of consumption and behavioural factors that have resulted in the observed transformation in the attitude of the UK grocery-shopping consumer.

De Kervenoael et al. (2006) suggest that an important influence on this shift was caused by the changing nature of the socioeconomic makeup of the UK. With an increasing number of single person households and growth of two-person households now commonly having two incomes, spare time has become a precious commodity (Baron et al., 2001, De Kervenoael et al., 2006, Hood et al., 2015). Trends suggest that consumers now spend less time cooking, are choosing to shop closer to home and to shop more frequently with smaller basket sizes (Dashwood, 2013b, Kenhove and De Wulf, 2000). Supporting research found that customers deciding to shop more than 3 times a week had increased from 9% in 1980 to 21% in 2002 (Hallsworth et al., 2010), rising again to 49% by 2011 (IGD, 2012). Declining basket size has, to some extent, been attributed to a decrease in both the average household size and the average number of children per women, therefore requiring reduced consumption levels (Dashwood, 2013b, De Kervenoael et al., 2006). Reductions in basket size are also linked to tighter spending following periods of financial downturn, with 59% of consumers concerned with price and choosing to buy less but shop more frequently in order to better monitor spending (IGD, 2012, ONS, 2015a).

This behaviour, although not limited to, is particularly seen in younger single consumer households, such as students or professionals who are characterised as time-poor, with low consumption levels and typically with poor cooking and storage facilities. Hallsworth et al. (2010) also suggest that this increased demand for convenience retailing is likely linked to a combination of general population changes in consumer behaviour. These included overall increases in the number of working hours and fewer statutory holidays provided to the working population, as well as more complex and irregular working patterns of consumers resulting in varied lifestyles within households (with spare time becoming less available and structured). There has also been a change and reduction of formal meal times, resulting in a rise in ready-meal consumption and an increase in the sale of fresh perishable products rather than dry or basic foodstuffs (Hallsworth et al., 2010). Hence it can be argued that these factors act as possible causes for the development and expansion of retailers into the convenience market.

These changing consumer trends have also been supported through research on observed consumer behaviour presented in Hallsworth et al. (2010). Their analysis of Portsmouth shoppers identified convenience as a key factor impacting decision-making during grocery shopping.

2.6 Evolving commerce channels

Developments in technologies have resulted in significant changes to the traditional shopping experience. New retail channels are available, offering new dimensions to retailing for both the retailer and consumer in terms of the fundamental shopping experiences. This section will reflect upon the evolution of commerce channels, considering the impacts to both the retailer and consumer as behaviours and retail experiences adapt and evolve.

2.6.1 Traditional channel – brick and mortar retailing

Traditionally sales in the retail industry operated in a physical form with business-to-customer interactions taking place face-to-face in physical locations such as stores, classified as 'brick-and-mortar' retailing (Otto and Chung, 2000). This still remains the dominant retail channel, with the highest proportion of revenue coming from sales in this sector (Birkin et al., 2002, Otto and Chung, 2000). The attraction of brick-and-mortar retailing can be attributed to the following characteristics and behavioural attitudes: by operating in a physical location, customers enjoy a social experience with others, and business transactions happen face-to-face, allowing physical human contact and personal experiences during the sales process (Maity and Dass, 2014). Consequently, when purchasing in stores, customers are able to experience products first hand. Inspecting and sampling products during the buying process enables consumers to make better informed decisions about which products to buy (Brynjolfsson et al., 2009, Otto and Chung, 2000). Along with the ability to undertake physical examinations, brick-and-mortar retail provides immediacy to the retailing process: customers receive the products at the point of purchase for use and are able to pay by a larger range of methods (Brynjolfsson et al., 2009, Otto and Chung, 2000). However, a number of disadvantages to brick-and-mortar retailing are also identified and predominately relate to supply-side issues, for example, costs and space (Otto and Chung, 2000). Retail space can be expensive and store size limits the level of products that a retailer is able to stock in comparison to the extensive product ranges that can be found online (Brynjolfsson et al., 2009). Presently much of the literature on retailing is focused on the developments of multi-channel retailing and how consumers and retailers interact and behave through multiple channels of commerce and, what role physical channels still play.

2.6.2 E-commerce

Many of the recent transformations in retailing, in both supply-side and consumer behaviour have been driven by developments and uptake in technology and the internet (Birkin et al., 2002, Kitchin, 2013, Sagiroglu and Sinanc, 2013). In the context of retail channels, this led to the development of e-commerce and, more recently, m-commerce (Elliott et al., 2012, Maity and Dass, 2014, Wrigley, 2009). E-commerce is the process of retailing via the internet using websites to browse, research and purchase products on a computer rather than the traditional methods associated with brick-and-mortar (Burt and Sparks, 2003a, Maity and Dass, 2014). Once products are purchased, suppliers fulfil the order delivering the product to an address given by the customer (Burt and Sparks, 2003a). The emergence of e-commerce in the late 1990s has provided a significant source of revenue in the retail sector with steadily increasing market shares and uptake by e-businesses (Brown and Dant, 2014, Wrigley, 2009). At the start of 2000 revenue generated from e-retailing in the US was $40billion (Bakos, 2001), in contrast, more recent figures now total $200 billion (Brown and Dant, 2014). This is clear evidence that e-commerce is becoming widely used as a retail channel, making up a sizable percentage of sales, suggesting a considerable change in consumer behaviour. The advantages of e-commerce include access to a much larger catalogues of products, providing customers with more choice and allowing retailers to target larger and new markets. E-commerce also removes catchment size restrictions. Retailers can operate at lower over-head costs, expanding their market without the need to build additional stores, thereby reaching a much larger proportion of the population. With the online retail experience, data is collected and stored about customer preferences, making it possible for retailers to offer personalised shopping experiences, suggesting related products or special deals to customers and developing a loyal retailer-customer relationship (Bakos, 2001, Brown and Dant, 2014, Otto and Chung, 2000). Disadvantages of e-commerce retailing include delivery costs in which either the customer or retailer absorbs the expense of delivery, and the lack of immediacy of the product: customers have to wait for the product and in some cases are not able to physically examine or try the product until it arrives.

A considerable change in regards to behaviour is that there are no longer temporal or spatial boundaries to shopping: the limits imposed by opening hours and geographical locations no longer apply. Customers are now able to purchase products on a global scale, seeking out items from international retailers, experiencing more products and seeking out better deals than found on the high street. E-commerce erodes the traditional shopping day with spatiotemporal factors no longer limiting trading and shopping hours. Laws limiting opening hours are irrelevant, allowing customers with access to the internet to shop 24/7, allowing retailers to sell at any time of the day and to any location (Burt and Sparks, 2003a, Brown and Dant, 2014), also thereby lifting the traditional spatial confines experienced in retailing, though the provision of services such as online grocery shopping. The power of prime locations and market dominance through acquisition of premium locations are of less importance, and as a result, it has been argued that the use and sense of place of locations may require reassessment in the age of internet retailing (Burt and Sparks, 2003a). Without the boundary of catchments and through extended product inventories, 24hr access and delivery to customers' homes, the purchase of goods can happen any time and from any connected [to the internet] device (Brown and Dant, 2014). E-commerce also opens up the market for new competitors that were previously not available to customers or have no physical presence in the market, such as the emergence of large online only retailers like Amazon and Ocado or access to more specialised stores offering products to a niche in the market (Burt and Sparks, 2003a, Noble et al., 2005), creating increased levels of competition.

2.6.3 M-commerce

Consumers and retailers alike are now making use of improved technological advances such as mobile phones, using them as part of the shopping experience (Deshmukh et al., 2013, Heitz-Spahn, 2013). Although fundamentally similar in its operation to e-commerce, m-commerce offers portability and increased accessibility opportunities through handheld devices and interfaces (Deshmukh et al., 2013, Maity and Dass, 2014). The increased accessibility has proved a popular channel with consumers and revenue opportunity for retailers. Levels of m-commerce in the US grew by 81% in 2012 alone (Wagner et al., 2013). This trend is likely to continue, as product penetration continues to grow with recent sales of smartphones exceeding the sales of PC devices (Wagner et al., 2013). In 2013 56% of the American population owned a smartphone and were able to connect via the internet (Einav et al., 2014), and a recent study suggested that in 2017 approximately 25% of all online retail transaction would take place on a mobile device (Turban et al., 2017). However, there are a number of disadvantages associated with m-commerce, such as screen size, particularly when dealing with information rich products and information-based products (Deshmukh et al., 2013, Maity and Dass, 2014).

2.6.4 S-commerce

S-commerce (social commerce) is defined as a new business model driven by social media sites whereby retailers facilitate the selling and purchasing of products via direct communication with customers, online/viral marketing and word-of-mouth. S-commerce is characterised by a unique set of characteristics, building on social experience and interactions of the customer and developing a relation-based market between customers and the retailer (Kim and Park, 2013, Zhou et al., 2013). Consumer interactions between themselves and the retailer using sites such as Facebook and Twitter, are some of the major drivers of s-commerce retailing. Retailers build trust in their services and the quality of their products through the social interactions of the consumers, through product reviews, sharing of posts and word-of-mouth experiences.

However, it is also possible for this to work negatively for retailers, for example when consumer experiences are not positive and can equally lead to widespread dissemination of negative publicity. Nevertheless relatively cost effective viral advertising and sharing of online blogs via s-commerce can indirectly drive sales and increase brand recognition driving interest in brands and products on a global scale (Zhou et al., 2013).

2.6.5 Omni-channel shopping and behaviour

Through the development of more competitive markets and technological improvements retailers have developed and responded to new multi-channel shopping behaviour by consumers. 'Omni-channel' retailing provides synergy to customer shopping experiences, across all retail channels through the creation of multi-channel strategies (Elliott et al., 2012). Omni-channel retailing allows customers to browse, research and purchase products with the ability to start the process of shopping using one channel, before seamlessly switching to another channel and finishing the process (Falk, 2014, Wagner et al., 2013). For example, several interfaces are now available with retailers making use of brick-and-mortar, e-commerce, m-commerce, virtual stores or employee tablets to engage with and connect customers to their service networks and products (Falk, 2014, Rowell, 2013). Using a combination of channels allows retailers to provide services to their customers, overcoming areas of weakness in one particular channel through the integration of a different channel for that part of the shopping activity (Wagner et al., 2013). If channels are operated and integrated equally with the full range of services and products marketed across each channel, this can lead to enhanced satisfaction in the purchasing processes, with retailers better able to meet the needs of customer preferences and drive further shopping related activities (Wagner et al., 2013). The impacts of omni-channel retailing are that customers have a wider range of services providing access to products in ways best suited to the needs of the customer at that moment in time, offering locational and temporal convenience. Consequently, the behaviour of the consumer utilising multiple channels within one, or across multiple retailers, incites the feeling of empowerment, regaining control of their shopping experience (Heitz-Spahn, 2013). In other words, consumers are able to tailor their shopping experience to best suit their own behaviour and individual needs (Brown and Dant, 2014).

2.6.6 Evolving commerce channels in the grocery sector

The developments in omni-channel retailing within the grocery sector have also had a considerable influence on consumer behaviour and the shopping experience. Take, for example, the development of home delivery: customers are able to order products online using a virtual shopping cart and have them delivered to their home at an appropriate and convenient time for them without having to physically set foot in a grocery store (Birkin et al., 2002). Within the

UK this was first offered on a national scale by Iceland and was closely followed by Tesco in 1996 (Birkin et al., 2002, Humby and Hunt, 2003). Since initially providing the service, Tesco was the largest online grocery retailer in the world and has a loyal online customer program with 380,000 monthly active users (Humby and Hunt, 2003). Asda unlike Tesco (who package orders at the nearest supermarket to offer home delivery) developed dedicated warehouses across the country to process orders and meet this demand. Their main goal behind this strategy, as opposed to Tesco's, was to increase levels of market penetration in areas where competitors hold the bulk of market share and Asda had low or no physical presence, gaining customers through economic and online driven behaviour (Birkin et al., 2002). An extension of this service is the concept of virtual stores, blending online shopping with a 'physical' store. Virtual stores are located in transport locations such as airports and train stations: customers use QR (quick response) codes next to photos of products, scanning items into their baskets using their mobiles and ordering the goods for home delivery at an appropriate time (Birkin et al., 2017, Rowell, 2013).

One of the issues grocery retailers face from offering home delivery is the relatively low profit margin gained from adding this channel to their portfolio of services. Possibly in response to these low profit margins, some retailers have developed an adapted business model, still making use of multiple channels in the format of 'Click and Collect' (Rowell, 2013). Rowell (2013) describes how customers using this service drive to the locations where the ordered goods are packed and ready for collection. Outlets of this kind, where customers 'collect' their groceries offer effective cost advantages saving money in distribution and operational costs, providing a lower cost strategy for geographical expansion than traditional store formats (Rowell, 2013). Likewise, retailers are able to package delivery services into single locations known as 'dark stores', which fulfil online orders and are not open to the public, helping to reduce costs. Consumers seeking this service demonstrate irregular and time-poor characteristics, ordering groceries via the internet to save time, but also being unable to set a delivery time, thereby picking the items up instead at their own convenience. It is plausible to assume a link with this tendency and the undertaking of multi-purpose trips, combining activities in order to make the most efficient use of time, e.g. combining shopping trips or as part of a journey or leisure activity. Retailers such as Tesco, who already have considerable national coverage, have adopted this technique adding the additional service to existing stores eliminating delivery costs. Furthermore, evidence suggests that customers who are purchasing through omni-channels such as 'Click and Collect' increase footfall in stores, driving additional sales, with customers buying 'extras' they may have missed or not purchased online, making this model significantly attractive for retailers (Bahn and Fischer, 2003). Uptake of online grocery shopping via e-commerce and omni-channel retailing has grown at a considerable rate in the UK market, with an increase of over 5% in the last five years with 26% of shoppers using online services regularly in January of 2015. Of this percentage, 11% of shoppers claimed to use online channels as the main shopping service with many shoppers also using a mixture of home delivery and click and collect options, switching to the most convenient channel (Henry, 2015). Omni-channel shopping using a combination of online and brick-and-mortar services and this behaviour is likely to continue to be exploited by supermarket retailers (Henry, 2015, Rowell, 2013).

2.7 Spatiotemporal determinants of consumer behaviour

Time availability and the utilisation of time by consumers for shopping can have considerable impact upon consumer behaviour as well as individual store performance. Shopping is a costly pursuit, requiring energy, money, information and time; and consumers are regarded as being as conscious of time as they are of the monetary controls when shopping and making purchases (Hornik, 1984). Thus, time has considerable impact on both the consumers as well as the retailers in terms of behaviour and trade. Time impacts consumer behaviour physically as well as psychologically; for example, through the time available, the time of day the shopping occurs, where a consumer is or a consumer's mood at the time of shopping (Solomon, 2013) and ultimately this greatly affects decision making during the shopping process. Time is a finite resource that is subsequently divided across activities according to each individual's needs at that point. Therefore, throughout the day consumers' behaviour, location, consumption patterns and needs will vary, exhibiting different patterns of consumption across space and time.

Evidence suggests that consumer perception of time has changed in recent years with consumers believing that they now, more than ever, have less free time available, a circumstance referred to as 'time poverty'. This belief has become commonplace with consumers and as noted in section 2.5.3, evidence clearly supports this belief (Hallsworth et al., 2010, Solomon, 2013). 'Lack of time' clearly has considerable impact on the behaviour, consumption and shopping experiences of consumers as well as the choice of shopping destination. It is also argued that this belief is in fact psychological rather than a physical change and instead is caused by the perception that less time is available, a view which is triggered by pressure caused by the increased number of options available in which to spend our time (Solomon, 2013). A behavioural response to this belief/problem is multi-purpose shopping (MPS): this is an attempt by consumers to optimise their time and has particularly been seen with consumers' grocery shopping, combining food shopping with other activities (Baker, 1996, Leszczyc et al., 2004). MPS is defined as the process of chaining shopping trips or purchases into one journey, either through visiting multiple stores during one trip or making multiple purchases from a range of products categories in one location (Baker, 1996, Leszczyc et al., 2004). The behaviour associated with MPS is inherently associated with consumers who feel they need to seek a better optimisation of their use of time while minimising the effort needed to complete tasks such as shopping (Baker, 1996, Leszczyc et al., 2004, Thill and Thomas, 1987). It is argued that multi-purpose shoppers will often bypass the nearest store in order to shop at an agglomerated site with multiple stores, choosing to travel further in return for the ability to combine trips and purchases within a single journey, consequently optimising time management and effort needed (Baker, 1996, Leszczyc et al., 2004). A second pattern of consumer behaviour, and not to be confused with MPS but in principle demonstrates a similar behaviour, is multi-purpose trip making (MPTM). MPTM is the process of undertaking multiple activities, of different purposes, at a number of destinations (Thill and Thomas, 1987, Timmermans et al., 2003). Consumers follow the same time saving tendencies joining activities into single journeys such as the school run, exercise, work, leisure as well as shopping (Birkin et al., 2010, O'Kelly, 1981, Timmermans et al., 2003). It is important to mention both MPS and MPTM as they both represented similar, yet separate lines of discussion and should not be assumed to represent the same process, a detailed discussion on MPS and MPTM can be seen in literature (Arentze et al., 2005, Baker, 1996, Leszczyc et al., 2004, O'Kelly, 1983, Timmermans et al., 2002). It is important to distinguish between MPS and MPTM because when modelling, while they are similar, they represent different patterns of behaviour and models simulating either behaviour should not be confused. Methods for modelling both MPS and MPTM are discussed in Chapter three. It is claimed that consumers who exhibit multi-purpose behaviour base decisions regarding store choice differently to single-purpose shoppers, assessing destination utility not only on an individual store's attributes or distance travelled to the store (but also on the utility of that store/site for continued shopping or activities from that destination) taking into account surrounding facilities (see section 3.4) and services (Timmermans et al., 2002, Leszczyc et al., 2004).

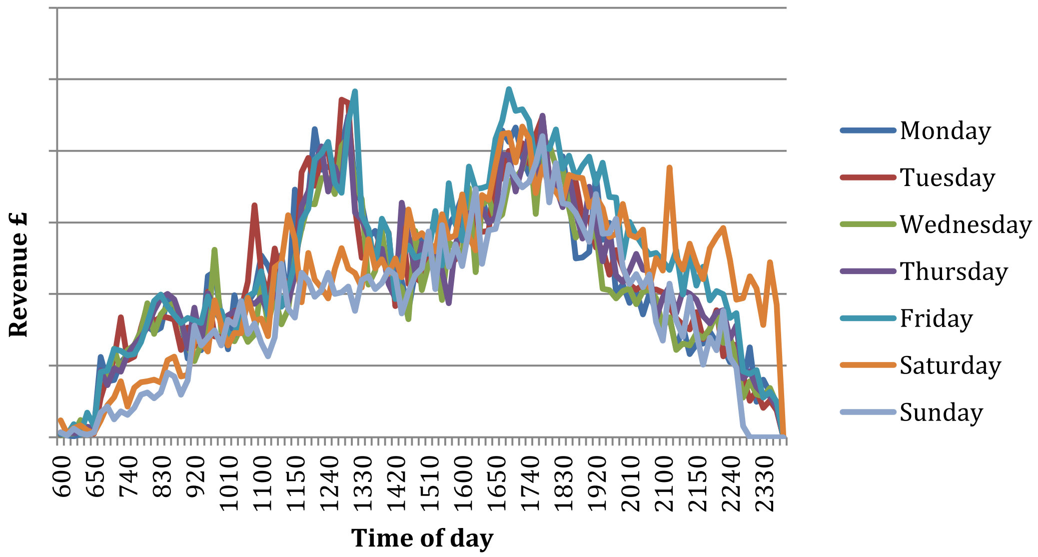
2.7.1 Spatiotemporal impacts on retail behaviour

The function of a particular space-time period has a considerable impact upon a consumer's behaviour at that given time and consequently individuals undertaking the same task will behave differently during individual time-space situations (Birkin et al., 2013). For instance, work-related time periods influence patterns of grocery shopping and the consumption of food and often occur in different locations to a consumer's home. Work-related consumer behaviour also varies considerably in terms of the patterns and needs of the consumer compared to a consumer shopping for leisure (Roy, 2004, Schwanen, 2004). This assumption that people behave differently in different spatiotemporal periods is further supported through a recent big data project. Analysis of crowd-sourced Twitter data found that distinct behavioural patterns existed at particular times and localities, e.g. work and home, following contextual analysis of tweets (Birkin et al., 2013, Birkin and Malleson, 2013, Malleson and Birkin, 2013a, Malleson and Birkin, 2013b, Malleson and Birkin, 2014). Traits associated with consumer behaviour during a particular space-time period, i.e. during work or returning home, were identified, demonstrating a spatiotemporally fluctuating demand.

While some acknowledgment that distinct space-time periods, and consequently distinct patterns of behaviour occur throughout the day has previously been noted (Birkin et al., 2017, Birkin et al., 2013, Roy, 2004), fewer articles have focused in detail on the specific characteristics of consumer behaviour and distribution of consumers over time and space and their impact on the grocery sector. Previous studies suggest there are several distinct time periods throughout the day in which grocery shopping and distinct consumer behaviours occur and these could be used to inform modelling techniques and provide insight into store performance. These periods represent discrete periods throughout the daily cycle and it is possible to argue that they demonstrate a clear relationship with food consumption and grocery shopping as well as demonstrating distinct patterns of shopping which are capable of impacting upon store performance (Schwanen, 2004, Syed Rakib-Uddin and Longley, 2014, Waddington et al., 2017). Evidence of this relationship, resulting from spatially and temporally fluctuating demand, is presented below in the form of sales profiles provided by a leading UK supermarket retailer which identifies increased revenue levels during commuting and working hours (shown in figure. 2.7). However the impact of spatiotemporal change and the impact upon store revenue is analysed in much greater detail in chapters four and five.

During the working day shopping activities have to be scheduled around work activities and unsurprisingly the amount of time allocated to shopping is limited by the amount of fixed time available, as well as by the location of work. Evidence suggests consumers who work longer hours will generally allocate fewer hours to shopping. Shopping behaviour during the day has been found to be most common during the two commuting periods, with a slight tendency for females to pursue shopping more frequently during these times and for longer periods than males (Schwanen, 2004). These shopping episodes (while representing the longest shopping period of the working day) more often than not tend to last less than any shopping periods originating from home. Consumers shopping during the commute period have also been noted to exhibit MPS behaviour, often combining shopping with travelling or other various activities, such as dropping children off or picking them up from school, buying petrol or a visit to a leisure activity.

Figure 2.7 - Typical weekly revenue profile for city centre convenience stores located within close proximity of densely populated workplace zones based on the store classification of convenience store types presented in Hood et al. (2015)



A rational assumption is that consumers shopping during commuting journeys will favour stores located on the route of travel or in close proximity to other activities that are undertaken at the time. Mode of transport also impacts consumer behaviour during these different time periods, such as whether a consumer shops with a car or is using public transport. Research has found that when the frequency of shopping episodes (or the number of activities during these times) increase, less time is allocated to each shopping trip (Schwanen, 2004). It is possible to suggest that consumers, particularly relating to grocery shopping, do so for necessity and convenience, attempting to avoid the utilisation of time outside of work and for leisure. This behaviour can reasonably be linked back to the notion of 'top-up' shops or 'food for now' and the decline of weekly shops, with consumers buying food more regularly and in smaller quantities. Consumers deciding to shop during the commuting period can have a considerable impact upon store performance and it has been noted that store performance will also vary considerably at these times due to location: e.g. convenience stores located at train, bus or at petrol stations often perform well during these hours. This impact has also been noted at a finer micro-geography level with consumers favouring stores located on the route that matches the direction of travel with stores on the 'correct' side performing better at that time of day (Birkin et al., 2002).

It is argued that the shortest temporal period allocated to shopping is during the middle of the work period, with a mean length of 18 minutes (Schwanen, 2004). It is claimed during the working day shopping is commonly associated with buying lunch, and due to time limitations and limited cooking facilities, consumers often demonstrate behaviour whereby they seek the closest convenient retailer offering ready to eat or 'food for now' products. A confidential industry report on shopping habits for consumer behaviour during work in an urbanised area found that 60% consumers chose to shop within a five-minute walk for their lunch (seen by the author only). Supporting anecdotal remarks by industry staff during correspondence noted similar patterns within their own customer behaviour. Therefore, it can be quite accurately reasoned that consumers shopping during the working day are highly time conscious, looking for a convenient shopping experience within close proximity where they can buy foodstuff for immediate consumption. Thus, any shopping behaviour is limited to a small catchment area surround the individual consumer's workplace. Consequently, a considerable amount of trade is made within supermarkets and convenience stores that are located in town/city centres or retail parks surrounding urban areas with high levels of workplace demand, which is likely to vary considerably over time as demand fluctuates.

2.8 Conclusions

This chapter has explored various supply and demand-side influences on consumer behaviour, providing a review of the types and patterns of behaviour as well as the evolution of the UK grocery market. The evidence demonstrates a diverse array of consumer types and supply side characteristics. This type of analysis, the disaggregation of consumers, is utilised by both commercial and academic researchers, offering insight into the type of behaviour that consumers exhibit. As noted, this review of consumer behaviour helped to demonstrate how behaviour in the current UK climate has evolved, and thus provided insight into the more recent patterns of consumer behaviour. Evidence on changes to the socio-economic and demographic makeup of UK consumers demonstrates the need to further understand the patterns of behaviour and how these can change over time.

A corpus of material has examined consumer behaviour theory and it is possible to infer the impact upon the retail network. However, limited research exists on applying spatiotemporal components of consumer behaviour in spatial interaction modelling with even fewer studies investigating the impact of spatiotemporal fluctuation on store revenue. The successful integration of the spatial interaction model with temporal extensions will broaden the scope of current models making it possible to model consumer behaviour across a much broader range of settings, improving the ability to reproduce more complex patterns of behaviour and temporal distribution of different novel demand types occurring in reality. Therefore, this current lack of research highlights an interesting and novel area of research. In the next chapter this thesis focuses on the literature relating to location based modelling and in particular the spatial interaction model used in the UK grocery sector.

Chapter 3 – Modelling consumer behaviour and a Review of Spatial interaction modelling

3.1 Introduction

This chapter will provide a review of retail location based modelling, paying particular attention to the modelling of store performance and consumer behaviour using spatial interaction models (SIMs). The aim is to also demonstrate evidence for the important advances and improvements to modelling that temporal components can provide. By providing a synopsis of related literature with a particular focus on the retail sector, as well as covering other modelling techniques commonly used in the retail sector, it attempts to present theoretical insight into modelling methods as well as discussion on approaches for incorporating different aspects of consumer behaviour. An exploration of temporal components (as well as the affects that modelling multiple demand layers within location modelling can have) will also be undertaken. Although only a limited amount of research incorporating temporal predictions has been undertaken in retail location analysis to date, the aim is that this will provide evidence supporting the need and rewards from temporal extensions to modelling (Newing et al., 2013a). A detailed analysis of the impacts upon revenue and observed temporal fluctuations is presented later in chapters four and five. Current literature relating to the application of temporal demand components in SIM is limited, and so comparable literature incorporating temporal-spatial modelling will also be considered (Martin, 2011b, Rasouli and Timmermans, 2013a, Timmermans et al., 2002), and this will attempt to provide some insight into existing methods of space-time modelling. The chapter will begin with an introduction to modelling within the retail industry, followed by the theory of SIMs and model development. The latter stage of this chapter will discuss current limitations and the benefits this research can provide, examining current methods for incorporating aspects of temporal consumer demand.

3.2 Location modelling in the retail industry

Retailing makes up a significant part of today's urban environment and consequently has been a keen topic of interest to researchers for some time (Dawson, 1980, Dawson, 2000, Pacione, 2009). After the early 1990s the study of retail geography began to focus heavily on economic geographies of retail, exploring issues around channel store-supplier relationships, retail space, corporate strategies and corporate culture (Lowe and Wrigley, 1996). There is no question that the study of retail geography is theoretically well developed and particular attention has been applied to the geographies of store location planning. The use of location modelling has become a well-established component of retail geography (Birkin et al., 2002). It is possible to address strategic issues, such as store performance, store location planning and the impact of competitors, by looking at how consumers interact across the retail environment i.e. the high street, or the home and how this impacts upon store performance (Birkin et al., 2002, Newing et al., 2013a, Pacione, 2009). Dawson (2000) p142 notes that "retail is a constantly evolving industry" and as such research and modelling in this field likewise need to evolve in order to remain relevant and provide solutions to the present research agenda. Recently there has been a rise in perceived value and uptake of big data analytics, and through the development of large volumes of data, it is believed that new opportunities exist for geographers to learn and observe the retail sector and consumer behaviour in much more detail. This has resulted in considerable change to the dynamics and understanding of the retail industry (Mint, 2014, PR Newswire, 2014, Birkin et al., 2017). This research aims to address the impact that spatiotemporal determinants have on store performance, providing novel insight and modelling developments.

Firstly, this section aims to provide insight into the use of models in retail industry, considering location modelling methods commonly adopted in the sector and the utility of such methods. Retail location analysis in its most simple form, site visits, has been used since the early 1900s, albeit in a much cruder and subjective manner than currently used today (Birkin et al., 1996, Sheppard and Plummer, 2009). Under the umbrella of the 'new retail geography' emerging in the 1990s, location analysis now incorporates detailed spatial analysis methods. Research has incorporated mathematical modelling, such as gravity models and entropy-maximizing models to establish expenditure flows and the revenue of different retail locations (Sheppard and Plummer, 2009). Outside of academia (Birkin et al., 2002), retailers often use 'gut feeling' and intuition for location planning based on site visits and experience to establish new stores instead of modelling (Reynolds and Wood, 2010, Wood and Reynolds, 2011a). However, Birkin et al. (1996) noted that since the 1960s the retail industry developed into a much more complex and sophisticated industry, creating an intensely competitive economy with continuous pressure for retailers to continually gain high profit margins.

Since the introduction of the revised PPG6 in 1996, it has become far more difficult for retailers to gain planning permission to develop large out-of-town developments, necessitating a change in expansion policies as well as a shift in consumer behaviour (Hood et al., 2015). This has led to a shift in strategy with the decision taken to develop new retail locations and retail channels previously considered too small and non-profitable; such as airports, train stations and traditional high streets (Birkin et al., 2002). Ironically locations that were previously withdrawn from or overlooked for development, such as the high streets and small towns, are now highly sought after investment opportunities, offering high market shares and satisfying new types of consumer demand (Birkin et al., 2002, Hood et al., 2015). Birkin et al. (2002) p245 write: *"we have been driven to find new and improved solutions* [through spatial modelling] *to business problems because of the value that these solutions provide"*. Retail organisations have become increasingly interested in the relationships between stores (supply) and consumers (demand) and the spatial interactions involved (Hernandez, 2007), in order to establish the best choices (i.e. new stores) or strategies to adopt (for existing stores) in order to drive businesses successfully and maximise profits, particularly when location decisions can be costly (Birkin et al., 1996). Hence, retail location analyses techniques, such as SIMs are now widely adopted (Cheng et al., 2007, Hernandez, 2007, Murad, 2003, Reynolds and Wood, 2010). The various methods used within the retail sector range from simple observation, sophisticated GIS analysis of spatial data to more complicated mathematical and statistical methods. Each approach has an array of applications. Table 3.1 provides a comprehensive synopsis of the methods used within the industry and is sourced from Wood and Reynolds (2011).

There are several perceived beneficial outcomes that make the processes of retail location analysis a valuable business strategy in the retail sector, suggesting reasons for why it has become common practice within the industry. Location analysis, for one, is a means for minimising risks in decision making, while maximising the potential for benefits through the ability to make better-informed decisions from the insights analysis provides. Likewise, the benefit of running such models is that they provide a relatively accurate prediction of revenue and store performance (Birkin et al., 2002). Retailers are able to test scenarios such as expansions of existing stores or the development of new store locations through these models. This helps them to understand the impacts on market shares, store turnover, customer profiles and behaviour, potential performance of competitor stores and to highlight areas in which they are currently underrepresented. Stakeholders are able to assess with confidence what is the most appropriate and profitable course of action to take (Birkin et al., 2002, Murad, 2003). Such analysis provides retail location analysts with a competitive advantage over those who do not utilise these techniques (Birkin et al., 1996, Hernandez, 2007, Wood and Reynolds, 2011a).

Table 3.1 - Location analysis tools used within the retail sector

Technique	Details	Indicative research literature
Experience/experimental	'Rule-of-thumb' procedures often employed 'on site' where the benefits of experience, observation, and intuition drive decision making.	Wood and Tasker (2008)
Checklist	Procedure to systematically evaluate the value of (and between) site(s) on the basis of a number of established variables.	Lilien and Kotler (1983)
Ratio	Assumes that, if a retailer has a given share of competing floorspace in an area, then it will achieve that same proportion of total sales available.	Rogers (1992)
Analogues	Existing store (or stores) similar to the site are compared to it to tailor turnover expectations.	Clarke et al (2003)
Cluster	Analysis of clusters in analogue store data to form groups and permit segmentation.	Schaffer and Green (1998)
Discriminant analysis	A screening tool or decision aid for low-value investments. Uses existing store performance to identify those variables that best explain the differences between preselected groups of stores. Site is then allocated to relevant turnover group.	Mendes and Themido (2004)
Multiple regression	Attempts to define a correlation between store sales and variables within the catchment that influence performance.	Morphet (1991)
Geographical information systems (GIS)	Spatial representation of geodemographic and retail data that is based on digitalised cartography and draws on relational databases.	Hernández (2007)
Spatial interaction/gravity/ entropy-maximising modelling	Derived from Newtonian laws of physics based on the relationship between store attractiveness and distance from consumers. May operate 'within' a GIS.	Birkin et al (2010)
Neural networks	Computer-based models explicitly represent the neural and synaptic activity of the biological brain.	Birkin et al (2002)

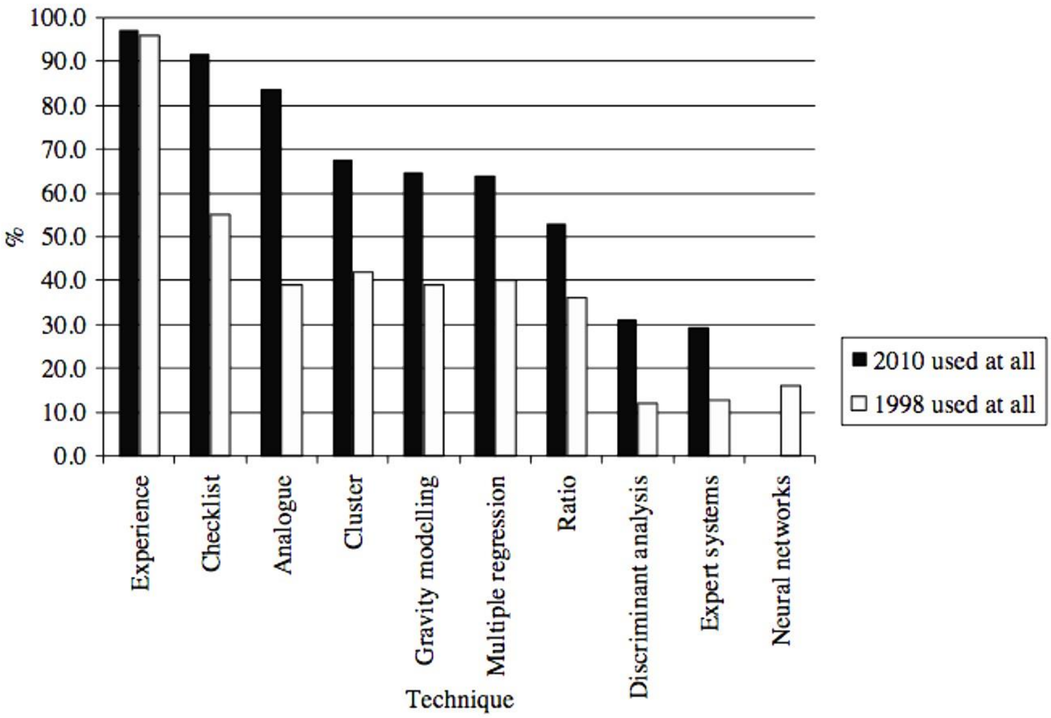
Source: (Wood and Reynolds, 2011a) p2470.

Methods of location planning are improving in utility and accuracy (Cheng et al., 2007, Hernandez, 2007) and it is clear that these methods are becoming more crucial in the decision making process (Birkin et al., 2017, Wrigley, 2009). Birkin et al. (1996) note some of the general benefits of commercial location which include increased turnover, more profit and better margins, with further detail presented in Birkin et al. (2017).

SIMs provide predictions of demand-supply interactions, thus, making them highly attractive techniques, hence the significant uptake in retail studies and in commercial teams.

Table 3.2 below, shows the level of usage for modelling techniques over a ten-year period. Model-based methods have experienced a significant increase in use, and while ‘experience’ remains the most common technique, all but one of the alternative methods (neural networks) used in location planning have increased (Reynolds and Wood, 2010). Research argues that the average uptake is also highly dependent on the type of location decisions being made: for example, rates were higher in relation to new store development analysis (Reynolds and Wood, 2010, Wood and Reynolds, 2011a). It can be argued that the increases identified are due to attempts to support more subjective location-based decision-making techniques with statistical observation and science by those retailers who have adopted more objective measures (Reynolds and Wood, 2010, Wood and Reynolds, 2011a). While SIMs represent but one of the modelling methods that have experienced an increased level of use (Reynolds and Wood, 2010), section 3.3.1 provides argument for adopting this technique when researching the grocery market.

Table 3.2 - Bar chart showing the level of usage for location-based modeling



Source: (Reynolds and Wood, 2010) p10.

3.3 The theory of Spatial Interaction

Since developments in the 1960s, SIMs have become one of the ‘go to’ methods for observing spatial interaction behaviour in many areas of research, such as commuting, shopping trips and migration (Huff, 1963, Patuelli and Arbia, 2013). Along with a distinguished history in the field of geography, there has been a considerable reliance on the SIM in studies of retail location analysis as well as in commercial use and it is an important and versatile tool for users, with a high degree of accuracy (Newing et al., 2014b).

The early models, commonly referred to as 'gravity models' derived their name through the concept of gravitational attraction, a Newtonian analogy: namely, that the gravitational force between two bodies is proportional to the mass of the two bodies and inversely proportional to the distance between them (Birkin et al., 1996, Birkin et al., 2002, Newing et al., 2014b). from the general gravity model came the 'law of retail gravitation' (Reilly, 1931) which in turn was the basis for Huff (1963) initial gravity model, expressed in equation 3.1 and forms the basis of SIMs used today.

The gravity model is written below in its conventional form as (adapted from Batty and Mackie (1972) and Wilson (1971)):

$$T_{ij} = kO_iD_jf(C_{ij}) \tag{3.1}$$

Where: T_{ij} is the interaction between origin zone i and destination zone j;

k is a constant of proportionality;

O_i and D_j are the amount of activity produced at origin zone i and attracted to destination zone j respectively ;

f is a generalised function of the spatial impendence between zones i and j: as f increases (C_{ij}) decreases, controlling the importance of distance; it is used for calibration;

(C_{ij}) is a measure of distance or cost between zones i and j.

3.3.1 Early modelling approaches: The entropy-maximising model

Particular problems of the gravity model, were inconsistencies in replicating known information; the model was aggregated with a poor forecasting capacity. For instance, a doubling of the origin population and destination attractiveness, results in a quadrupling of flows, rather than doubling the interaction as expected (Senior, 1979, Thompson, 2013). Subsequently, the family of models proposed by Wilson (1971) were designed to overcome these issues through the introduction of constraints in the system. Adding constraints within the model helps to improve the model's capabilities, overcoming many of the early criticisms of the model surrounding the aggregate nature of the approach and poor predictive capabilities. Wilson's family of models include constraints at a macro-level, maximising the number of unbiased flows at a micro-level, yet remaining consistent with the aggregate constraints, applied with an equiprobable interaction (Roy and Thill, 2004, Senior, 1979). Within the entropy-maximising SIM family, four cases exist which lend themselves to the modelling of different spatial interaction phenomena through the incorporation of different constraints on the SIM framework (Wilson, 1971).

The four models purposed were:

1. Where neither O_i or D_j are known, this represents the unconstrained case;
2. Where O_i is known, this represents the production-constrained case;
3. Where D_j is known, this represents the attraction-constrained cases;
4. Where both O_i and D_j are known, this represents the doubly-constrained cases.

For the purpose of this research the modelling approach adopted from Wilson's SIM family is the *production-constrained model*.

The model has two main hypotheses:

(1) Flows between an origin and destination will be proportional to the relative attractiveness of that destination *vis a vis* all other competing destinations.
(2) Flows between an origin and destination will be proportional to the relative accessibility of that destination vis a vis all other competing destinations.

Source: (Birkin et al., 2002)p 152.

The production-constrained model is regarded in the literature as the most commonly adopted model when used in grocery retail applications (Newing et al., 2014b). However, in reality, real world applications of SIMs require significant levels of customisation in order to more accurately predict sales (Birkin et al., 2002), which has been demonstrated through recent examples of SIM applications with significant customisation (Birkin et al., 2010, Newing et al., 2014b). The aggregate production-constrained model is expressed in the follow equation (3.2):

$$S_{ij} = A_i \, O_i \, W_j^{\alpha} exp^{-\beta c_{ij}} \tag{3.2}$$

where,

S_{ij} is the interaction of people or flow of expenditure between origin i and store j.

O_i represents the demand or amount of expenditure available in origin *i*: this is often derived from the combination of average weekly spending by some household classifications, i.e:

$$O_i = \sum_g H_i^g F^g$$

(3.3)

where,

H_i^g is the total number of households in origin i by type g and F^g is the average weekly spend on groceries by that household type g.

W_j^α is the measure of attractiveness of store j, α is a balancing factor influencing the importance of the attractiveness variable for store j.

$exp^{-\beta C_{ij}}$ is the distance deterrence term indicating the propensity to travel, integrating β, the distance decay parameter , and controls flows by influencing the importance of distance. C_{ij} accounts for the cost incurred i.e. the distance or travel time between origin i and store j.

A_i is a balancing factor that ensures that all demand in the area is allocated to centres within the model, which is shown in equation (3.4).

$$A_i = \frac{1}{\sum_j W_j \times exp^{-\beta C_{ij}}}$$

(3.4)

ensuring that:

$$O_i = \sum_j S_{ij}$$

(3.5)

The SIM is widely regarded to be reliable and accurate (Birkin et al., 1996, Birkin et al., 2002, Wilson, 2000). Justification of the use of modelling technique in relation to this thesis was due to a combination of factors, influenced by literature, data and in particular current industry practice. In a recent study on location analyst teams within the retail industry by Reynolds and Wood (2010), explored the use of store performance and site assessment techniques. In particular, grocery retailers displayed a tendency to favour a subset of the analytical methods: analogue, multiple regression and SIM. The widespread use of SIMs (often the preferred method) further justifies the utilisation of the SIM technique within this research for studies on the grocery sector (Reynolds and Wood, 2010). The fact that the production-constrained SIM is not only widely utilised but is also arguably the most appropriate model choice for grocery store revenue estimation is widely shared throughout the corpus of location modelling literature (Newing et al., 2014b, Senior, 1979, Thompson, 2013, Wilson, 1971).

The literature highlights the high levels of achievable accuracy (often within 10% of reality) of SIMs, further demonstrating the appropriateness of this model choice (Birkin et al., 2002, Newing et al., 2014b). Newing et al. (2013, 2014) showed that it was possible to further improve model predictions for grocery store revenue estimation using temporal seasonal components and multiple demand types, including different consumer behaviour at different times of the year. The incorporation of seasonal demand and visitor behaviour boosted prediction rates in some cases to within 5% of actual revenues. This research has comparable aims and therefore the successful implementation identified in Newing et al. (2014) in modelling spatiotemporal demand variants demonstrates the attainable and also beneficial qualities of this research.

In the grocery industry, retail location modelling is a valuable resource and becoming more commonplace, providing answers to spatial problems and queries in a commercial and academic context (Birkin et al., 2002, Cheng et al., 2007, Hernandez, 2007, Nakaya et al., 2007, O'Sullivan and Perry, 2013). It is also noted that for modelling outputs to be affective commercially and at the corporate strategy level, it is important that information provided through location based modelling is suited to the needs of the organisation and comprehension of management staff (Wood and Reynolds, 2011a). Therefore adopting a method that is already widely used, with the analysis and data generated, already commonly understood further strengthens the use of SIMs and improving its capabilities from temporal and demand-side disaggregation (see later chapters).

3.3.2 Model calibration

Within the grocery sector an accepted accuracy threshold of +/-10% for revenue predictions is considered as the industry standard (Newing et al., 2014b). In order to attain this high degree of accuracy it is necessary to model predicted consumer flows as closely to observed consumer flows, thereby optimising the model conditions to replicate actual consumer flows within the grocery sector (Birkin et al., 2010, Newing et al., 2014b). SIMs can be calibrated through both the demand-side and supply-side parameters (Alpha and Beta), effectively replicating consumer behaviour in relation to store choice and travel behaviour (Birkin et al., 2010, Newing et al., 2014b). The effectiveness of calibration measures is primarily linked to the availability of observed data whereupon appropriate goodness-of-fit (GOF) statistics are used to measure model performance, with the calibration process identifying the best parameter values. Traditionally access to detailed consumer flow data has been limited in the past, either through lack of access or conversely, when data is available the scope or representation of the data are limited (Birkin et al., 2010, Newing et al., 2014b).

That said, calibration of the model using observed data is a crucial stage in the development of a SIM. It is important because ensuring that the model recreates observed conditions indicates that the model is working as expected and accuracy can be demonstrated. The absence of observed interactions reduces the overall level of confidence in a model's performance and limits the models capability (Shepherd and Thomas, 1980).

To date a number of reviews have detailed the calibration and GOF statistics most commonly used (Batty and Mackie, 1972, Diplock and Openshaw, 1996, Guy, 1991, Openshaw, 1977, Harland, 2008). Again focus will be placed upon the production-constrained SIM and the methods commonly adopted for calibration and validation. A commonly adopted statistical approach for the calibration of SIMs, is the use of observed and predicted average trip distance (ATD) (Birkin et al., 2010, Guy, 1991, Thompson, 2013). ATD represents the average distance an individual will travel i.e. from their home to a store and is expressed below.

Predicted ATD (ATD^{pre}) can be written as:

$$ATD^{pre} = \frac{\sum_{ij} S_{ij} C_{ij}}{\sum_{ij} S_{ij}} \tag{3.6}$$

where,

S_{ij} represents predicted flows between i and j.

and observed ATD (ATD^{obs}) can be written as:

$$ATD^{obs} = \frac{\sum_{ij} \check{S}_{ij} C_{ij}}{\sum_{ij} \check{S}_{ij}} \tag{3.7}$$

where,

$\check{S}_{ij}$ represents observed flows between i and j.

The exercise then is to minimise the difference between the two, preferably so that:

$$ATD^{o} = \frac{ATD^{pre}}{ATD^{obs}} \tag{3.8}$$

where,

ATD^{o} represents optimal model conditions (post-calibration) in which predicted and observed ATD have been balanced through the identification of an optimal parameter value.

While commonly used, a criticism of this approach is that ATD is potentially one dimensional, in the sense that positive and negative values can be balanced out in the calibration process without accurately matching behaviour. It is also possible for several combinations to produce the same result and this risk is even higher when performed on an aggregate population. An alternative means of GOF statistical validation of model performance is the Standardised Root Mean Squared Error (SRMSE), where the square root of the sum of all errors squared is divided by the matrix dimensions, with the subsequent value divided by the average interaction value for standardisation. SRMSE values have a lower limit of 0 indicating a perfect prediction and a variable upper limit dependent on the distribution of observed flows, although this is typically 1, and indicates a poor prediction necessitating further calibration of parameter values. SRMSE is calculated as follows:

$$SRMSE = \frac{\left[\sum_i \sum_j (\check{S}_{ij} - S_{ij})^2 / m * n\right]^{0.5}}{\left[\sum_i \sum_j \check{S}_{ij} / m * n\right]} \tag{3.9}$$

where,

$\check{S}_{ij}$ represents observed flows between i and j,

S_{ij} represents predicted flows between i and j, and

*(m * n)* are the dimensions of the interaction matrix.

A third approach is R^2, and in this case a value range of 0 – 1 exists. The closer the R^2 value is to 1, the greater the link between predicted and observed flow data i.e. if a value of 1 is achieved this indicates an exact correspondence between the two data. A value of zero indicates the opposite, reflecting no relationship.

R^2 is calculated as follows:

$$R^2 = \left[\frac{\sum_i \sum_j (\check{S}_{ij} - \bar{S}_o)(S_{ij} - \bar{S}_p)}{\left[\sum_i \sum_j \left(\check{S}_{ij} - \bar{S}_o\right)^2 * \sum_i \sum_j \left(S_{ij} - \bar{S}_p\right)^2\right]^{0.5}} \right]^2 \tag{3.10}$$

where,

$\bar{S}_o$ represents the mean of $\check{S}_{ij}'s$, and

$\bar{S}_p$ represents the mean of $S_{ij}'s$.

For the purpose of this research, ATD will be used for the calibration of the beta parameter, and R^2 and SRMSE will be used for the assessment of model performance and statistical validation of model calibration. R^2 and SRMSE have been noted to outperform other methods available and will be used for GOF statistics validation to improve the model accuracy (Dennett, 2010).

It is also important to consider Birkin et al. (2010) findings, noting that to avoid over-parameterisation of a model and artificially fitting the model to the observed data, it is important to use a control sample, testing the capability of the model and proving that the model works on similarly performing stores. One such solution is to withhold some observed data: fundamentally, the model is then calibrated using only a partial data source so that it is then possible to test the model predictions (after calibration) on the withheld observed data (or control group). For the purpose of this research, calibration will therefore be performed focusing on a partial sample of stores using observed flows, with the remaining observed flows for stores with sufficient data withheld for testing.

3.3.3 Model limitations

While SIMs have been proven to be an appropriate method of prediction of store revenue and consumer behaviour they are not without limitations. While models tend to be reliable and operate within a high level of accuracy it would be naïve to refrain from any acknowledgment of the limitations and potential criticism. Common geographical problems that affect SIMs, although not limited to modelling of this type and which likewise represent issues for spatial analysts generally, are the ecological fallacy and Modifiable Areal Unit Problem (MAUP). In this research attempts were made to limit the effects of both of these problems. Firstly, the impact of MAUP is approached with the use of population weighted centroid data in an attempt to make point based data, e.g. population within districts, more representative of actual distributions as well as the use of smaller areas such as LSOAs over MSOAs and LADs. Secondly, ecological fallacy is controlled to some extent through disaggregation of the model parameters (presented in chapter six), incorporating different consumer behaviour within the population, such as applying different β and α values to account for different types of consumer, (in this instance accounting for their output area classifications (OAC), and at the very least this will reproduce differences in consumer behaviour and store choice (Newing et al., 2014b)). Boundary effects are also of concern when using a SIM enclosed within a physical boundary. No knowledge of expenditure outside of the study area is known and therefore flows originating from and distributing to the 'excluded' zones are unknown, often resulting in unrealistic flow estimates along the edge of SIMs. In an attempt to eliminate this issue, Birkin et al. (2010), proposed the use of a boundary-free model, which demonstrated promising results. However, the incorporation of the boundary-free approach in this research is beyond the scope and resources of this project.

Data problems are also an issue in SIMs, although this is somewhat common in all commercial retailing research. Calibration of SIMs using observed data, as previously noted, is crucial to the model development and production of an accurate system. However, even if access is granted to commercial data for a specific retailer, it is possible to generate bias within the model through over parameterisation of the model variables to a specific retailer (Birkin et al., 2010, Lovelace et al., 2016). Birkin et al. (2010) also note several other problems related to data issues in the construction of SIMs. Demand estimation can be problematic: data may not be routinely collected, representative of the aggregate population or a representative sample size, which can lead to errors occurring. The methods of determining store attractiveness for supply estimates also present limitation issues effecting SIM reliability. Traditionally the measure of store attractiveness has been derived from store size. However, other store characteristics, such as brand, price, parking facilities, in-store services and surrounding facilities and services are also known to be important attractiveness components that can be included, which are likely to be of importance at certain times of day. Birkin et al. (2010) offer further explanation of a complex attractiveness W_j variable.

Another potential criticism is that a production-constrained SIM, due to the constraints of the equation, distributes all available flows from each origin as a single event with all possible interactions happening simultaneously, with all flows terminating at each available destination (Rasouli and Timmermans, 2013a). In other words, shopping behaviour is assumed to occur as a singular event in time without any consideration of past or future events, even if this is not the case. The assumption that all grocery shopping behaviour is single purpose is limiting and modelling of store performance accounting for temporal components of grocery shopping such as MPTM is hoped to improve the models. Research has suggested that 74% of non-grocery and 63% of grocery shopping trips are multi-purpose (O'Kelly, 1981). Therefore, a typical SIM's limitations in accounting for multi-purpose (MP) behaviour may result in inaccurate representation of flows and similarly a poor reproduction of actual consumer behaviour in these circumstances. However, it is possible to argue that in some cases grocery shopping still has a tendency to be conducted as a single-purpose shop due to the inherent nature of food shopping: items are perishable and bulky and so it is likely that trips will be conducted closer to home as a single event. However, by accounting for additional demand types and at different times of day it is possible to simulate the impact of new types of spending that are linked to different activities such as work, tourism or education. The important additional disaggregations included are discussed in the remainder of this thesis.

3.4 Alternative modelling approaches

Why is time important to a SIM? Recent studies into MP behaviour indicate that a number of points in space and time exist, acting as sources of trip-making and patterns of different behaviour throughout the urban environment (Malleson and Birkin, 2013b, Malleson and Birkin, 2014). These authors noted key locations or times, termed 'anchor points,' within observed patterns of behaviour and it became apparent that specific types of behaviour and movements occurred at these points. While the research was focused on the development of space-time modelling at an individual level, the patterns of behaviour and trip-making were identifiable at a more aggregate level with patterns observably associated with particular space-time periods, e.g. home, work, education and leisure (Malleson and Birkin, 2013b, Malleson and Birkin, 2014). Utilising available data sources appropriate for modelling, it is possible to replicate these differing periods of temporal behaviour within the SIM approach. Consistent evidence supporting the assumption that behaviour and patterns of movement change, but are distinguishable, within different space-time periods was noted previously in chapter two. This is demonstrated further in the context of grocery retailing in chapters four and five, and forms the foundations of this theses' novel contributions to academia. Therefore, incorporating temporal components of behaviour based on the assumption that behaviour will change through space and time (i.e. individuals at work will behave differently from when they are at home) affords this research the opportunity to represent more complex and temporal patterns of behaviour occurring in reality.

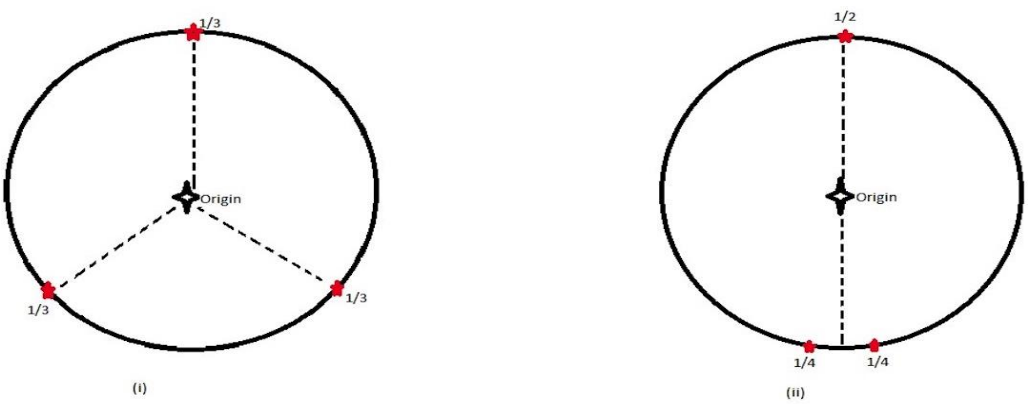
Away from the field of SIMs, alternative approaches in modelling temporal and multi-purpose behaviour include the use of Markov Chain Models (O'Kelly, 1981) and the household-level Multipurpose Shopping Model (Mulligan, 1983, Mulligan, 1987). The former, being the work of O'Kelly (1981), adopts Markov chain theory. Trips are defined as home-based, originating and terminating at home and works via the following assumptions: namely, that each stop is made for one of a given range of purposes, i.e. work, shopping, recreation. After every stop there is a certain probability of continuing to another location, and purpose, or of ending the trip. At each stop the probability to continue to another location or purpose does not necessarily remain constant, changing at each destination. At the beginning of the activity the distribution of location-purpose combinations for the first destination is known and the probability determining future behaviour is dependent on the present situation and does not account for previous actions (Thill and Thomas, 1987). The individual is then left to move through the system until they reach the final destination. The latter work, Mulligan (1987), works via the assumption that consumers will seek the nearest and least costly destination capable of fulfilling the demand for a particular good (Thill and Thomas, 1987). In this case this is based on the cost of the good at the destination and the cost incurred to travel to that destination and assumes that the destination decision takes into account the order of goods being purchased. A consumer will then choose the destination with the lowest cost involved for purchasing the required goods. Upon reaching this destination a decision is then made as to where the next good can be purchased for the least cost incurred. The model takes into account the agglomeration of stores, meaning that in the model customers are able to purchase all items in a single location, if the cost of purchasing the items in one location is cheaper than the costs incurred to travel and purchase these at a cheaper destination.

While these approaches are capable of generating the interactions of MPS flows at different times, the models are not only complex, requiring high levels of computation, but further issues prevent the use of these approaches. First, the models account for the interaction flows through trip frequency or number of stops at destinations as opposed to predicting monetary flows at a store as in a SIM. While theoretically it could be possible to revise models so that trip making and spending were predicted with unlimited access to data detailing observed patterns of all behaviour and spending, without full access to a complete universal dataset at an individual level, it would not be possible to build an MPTM model capable of predicting all grocery/non-grocery shopping and non-shopping activities accurately. Secondly, without incorporating all possible destinations such as non-food shopping or pursuits in addition to grocery destinations within a MPTM model, it would be difficult to predict trip-making behaviour, as there is no complete data available to calibrate or validate these kinds of models. It is also noted that in these models, due to severe computational problems, sample size is often restricted to smaller arrays and while a larger size is theoretically possible, even then it is argued that an array of 180 areas would encounter considerable problems utilising this approach (O'Kelly, 1981). Since the SIM being developed in this thesis far exceeds that capacity, incorporating 1300+ grocery stores and over 1000+ LSOAs it is unlikely that the modelling approaches discussed here would represent and alternative solution.

Alternatively, spatiotemporal behaviour such as MPS has been modelled using discrete-choice and computational agent-based models (Arentze et al., 2005, Rasouli and Timmermans, 2013a, Rasouli and Timmermans, 2013b, Thill and Thomas, 1987, Timmermans et al., 2002). Discrete-choice modelling is based on utility-maximising theory: this assumes the choices are based on a series of attributes about an individual destination. It is assumed that individuals then derive a level of utility according to the attribute levels about locations, combining destinations together into a trip with an overall measure of utility based on the total attributes. Available shopping trips are assessed individually, with the trip that maximises utility being chosen as the appropriate action (Thill and Thomas, 1987, Timmermans et al., 2002). In other words, according to the purpose of the shopping trip, shopping destinations are assessed for their convenience and destinations are in turn considered in a series of alternate available trip choices. The shopping trip that offers the best efficiency to the individual is then chosen. Agent based models work on the assumption that individuals will not necessarily choose the optimal path, instead relying on context-dependent heuristics to influence MP behaviour (Timmermans et al., 2002). Rasouli and Timmermans (2013a) note how this era of modelling approaches suggest that shopping is part of a household's and individual's daily life. Agent-based models imply that shopping is considered probabilistic in nature and consequently models allow MPTM and MPS behaviour as well as reflecting temporal dimensions such as working hours and store opening times, subsequently, simulating space-time behaviour and exhibiting more complex patterns of consumer behaviour. A relevant example of agent-based modelling is the Albatross model system (Arentze and Timmermans, 2000). This was originally developed to simulate daily activity patterns and though the model is not a dedicated shopping behaviour model, shopping represents one of the activity types included in the model making it possible to simulate MPTM and MPS behaviour with regards to trip frequency (Rasouli and Timmermans, 2013a, Rasouli and Timmermans, 2013b, Timmermans et al., 2002). As noted, the models are capable of simulating MPTM and MPS type behaviour; activity choices are derived through a series of constraints: situational, institutional, household and spatial-temporal constraints which when considered together impact on activity choice. The individual then chooses feasible activities based on the current combination of constraints (Arentze and Timmermans, 2000). The difference between this approach and others is that the rules which dictate behaviour choice are constantly evolving with the individual learning heuristically in the model environment and adapting until the best result is found (Arentze and Timmermans, 2000). In some instances the individual may choose a number of locations that correspond with a particular outcome. In other words, even though a location may initially seem inferior to another, due to its connection with a particular outcome, it becomes the best choice for that given shopping purpose. Additional detail of this approach can be seen in (Arentze and Timmermans, 2000, Rasouli and Timmermans, 2013b, Timmermans et al., 2002). This approach requires large scale consumer surveys to determine the individuals utilities, which this research does not have access to, furthermore when aggregated these models get to be similar to disaggregated SIMs. Therefore, these models will not be used for this research.

The work of Fotheringham in the early 1980s (Fotheringham, 1983, 1984, 1986) argued that the basic SIM failed to account for the distribution of destinations in relation to one and other, leading to a misspecification of reality and the subsequent development of the competing destination model (Fotheringham, 1984, Newing et al., 2014b, Roy, 2004, Roy and Thill, 2004). In a conventional SIM, if two destinations of equal accessibility and attractiveness were available, the model would assign a share of [½, ½]. If a third destination, *ceteris paribus*, was added in, as shown in figure 3.1(i), the model would assign a share of [⅓, ⅓, ⅓]. Forthingham noted that if the third destination were in fact located within close proximity to either the first or second, then the conventional model would assign shares of [½, ¼, ¼]; see figure 3.1(ii) (Fotheringham, 1983, Roy, 2004, Roy and Thill, 2004). It is unlikely that this distribution of flows would occur in reality. Instead it is more likely that the stores within close proximity would do better, with individuals attracted by the MP opportunities offered by the clustered destinations. For example, consumers are not only attracted to a particular store, but also by the fact it is close to adjacent stores. Each individual store increases the overall attractiveness of a destination where stores cluster, implying that consumers are likely to undertake MP behaviour because sites with multiple store are perceived as more attractive (Fotheringham, 1983, Fotheringham, 1984, Roy, 2004, Roy and Thill, 2004). This concept was noted in the previous chapter, which discussed consumer behaviour (section 2.8.1) noting that consumer destination choices in MP behaviour were influenced by the surrounding facilities and opportunities available from that destination and is reiterated here. Consequently, when stores cluster together they benefit from the agglomeration effect of consumer behaviour, experiencing higher levels of flows which are not accounted for in the conventional SIM. The attractiveness of destinations where stores are clustered together will increase as more stores are added into the group until a point where the attractiveness plateaus, irrespective of additional stores (Fotheringham, 1983, Baker, 1996).

Figure 3.1 - Retail destination configurations adapted from (Roy, 2004) p42, illustrating the distribution of market shares for stores and misspecification of a conventional SIM.



Adapted from Fotheringham (1983,1986) and Roy and Thill (2004) the competing destination model is expressed as follows:

$$S_{ij} = O_i W_j^{\alpha} A_j^{\gamma} \exp - \beta_{C_{ij}} / \left[\sum_j W_j^{\alpha} A_j^{\gamma} \exp - \beta_{C_{ij}} \right]$$

(3.11)

Where;

S_{ij} represents flows between origin i and destination j, O_i represents trip origins, W_j is floor space of destination j and C_{ij} are home-based travel costs between origin i and destination j. α is a scaling factor of floorspace at W_j and β is the distance decay parameter.

with the potential measure of competing destination A_j defined as,

$$A_j = \sum_{l \neq j} W_l \, c_{jl}^{-\beta} d_{jl}$$

(3.12)

where;

W_l is floor space of destination l, c_{jl} represents the unit cost for travel between primary destination j and nearby destination l. The binary variable d_{jl} will be one if destination l is within a predefined range of destination j and will be zero if outside this range. The scaling index γ, part of the competing destination potential term A_j^{γ}, indicates the cluster effects: a positive value demonstrates the presence of agglomerative effects and a negative value implies the presence of competition effects.

The positive response of agglomeration effects upon the grocery market has less of an impact. Evidence suggest that supermarkets tend not to exhibit agglomerative behaviour due to consumers gaining little from comparison shopping for groceries (Krider and Putler, 2013) cited in Newing et al. (2014b). Consequently, grocery stores within close proximity to one another are more likely to compete for sales (which Fotheringham noted as the competition effect), resulting in reduced levels of interaction between stores (the opposite of the agglomeration effect), with stores in relative isolation potentially capturing a larger market than when stores are clustered together (Fotheringham, 1983). In this instance, the competition effect is already being accounted for in a SIM and will therefore be used instead of Fotheringham's model.

Extensive disaggregation of demand type is an alternative approach for incorporating various spatiotemporal fluctuations in a SIM and thereby improving the reliability of models (Birkin et al., 2010, Hernandez, 2007, Newing et al., 2014b): for example, the incorporation of additional demand types such as workplace demand within a model. If the usual place of residence is known, and the usual place of work is known for individuals and this data is incorporated into a SIM as an additional layer, it is possible to simulate MP behaviour. In this instance, the place of work would represent the primary destination from their residence and secondary destinations would be stores clustered in the vicinity of the workplace. Following similar theory to the competing destination model, stores would need to be within close proximity to the workplace to be perceived as possible secondary destinations. This could be controlled through beta, using a high distance decay value, reducing the likelihood for workplace demand type to travel too far, ensuring realistic behaviour. As noted in the previous chapter, as well as supported by Birkin et al. (2010), MPTM behaviour such as home-work-shop-work-home has previously been observed and for the workplace population the distance travelled from work is typically minimal (see section 2.8.1). Utilising data on different demand types, particularly spatiotemporally structured data, it may be possible to build a series of detailed demand layers simulating the flow of people throughout the SIM study area. This could provide realistic population movements generating time specific revenue predictions, presenting a much more realistic picture of reality and consumer behaviour. An example of a similar technique being implemented incorporated a spatiotemporally informed tourism demand layer, which accounted for additional revenue and different consumer behaviour occurring in different locations and at certain times from seasonal visitors (Newing et al., 2013a, Newing et al., 2013b, Newing et al., 2014a, Newing et al., 2014b). Spatial-temporal data offer potential novel and insightful opportunities for SIM predictive capabilities particularly when incorporated within commercial data. However, few SIMs outside of commercial teams are fortunate to calibrate from such extensive and detail commercial data. Therefore, the level of detailed data provided for this research offers an exciting opportunity for considerable development within academic literature.

3.5 Conclusion

Following the discussion of consumer behaviour in the previous chapter, this chapter examined the methods used to predict store performance and consumer behaviour, detailing the conventional SIM and other potential models. This chapter has sought to explore location modelling within the grocery sector focusing on SIMs in an attempt to understand the utilisation and demonstrate the potential opportunities for temporal refinement for the SIM. The review of existing literature, while demonstrating that there is not only a need to further improve the applicability of the SIM through the incorporation of temporal aspects of consumer behaviour enabling an improved ability to model store performance, but that this is well within the scope of this research. The intended refinements to the model not only provide a novel development to SIMs but also supplement the existing knowledge and applications available. The SIM approach is not without limitation and it is not the only location modelling method available to retail. However, the discussion clearly illustrated the widespread utilisation of the model and that the intended developments this project seeks to achieve are suited towards the grocery industry and address current gaps in academia.

While evidence of many theoretical extensions to SIMs have been acknowledged by others (Birkin et al., 2002, Newing et al., 2014b), there are fewer examples detailing the type of extensions to improve actual model performance and applicability in the real world using spatiotemporal data (Newing et al., 2014b). This can be argued to be partly due to the lack of published work, either by consultants in the commercial sector [who have little incentive to publish client reports] or due to restrictions put in place on academics by commercial partners, thereby increasing the need for such work to be established. A number of opportunities have been identified in the limited discussion available on this subject. Although currently lacking academic investigation or commercial application it is maintained that if addressed (and SIMs are refined incorporating spatiotemporal behaviour) they have considerable potential for improving model performance and insight for both commercial and academic partners (Birkin et al., 2010, Newing et al., 2013a, Newing et al., 2013b, Newing et al., 2014b).

Demand side disaggregation currently remains uncommon (Newing et al., 2014b), and the modelling of multiple demand layers and behavioural changes, e.g. residential, work, leisure and how these vary throughout the day (Birkin et al., 2010, Newing et al., 2013a, Newing et al., 2014b), has been limited. Research reflects this view and it is argued that temporal changes have a substantial influence on market conditions or how individuals choose to behave (Birkin et al., 2013, Newing et al., 2013b), as have the impact of specific demand types on behaviour, shopping and trade patterns (Birkin et al., 2010). Therefore, they represent important factors to address when undertaking modelling techniques (Birkin et al., 2013). Consequently, the incorporation of temporal demand layer extensions, such as tourism demand across a temporal cycle, into models has the potential to improve the accuracy of predictions considerably (Newing et al., 2013a, Newing et al., 2013b, Newing et al., 2014b). If temporal components continue to be excluded, it may repeatedly lead to underestimations in revenue for certain stores particularly within the convenience store format. The impact of time is yet to be fully explored and while some retailers are beginning to examine short-term changes in demand and the impact at a store level (for instance local events or the weather) there has been little attempt to incorporate these at a more strategic level for revenue predictions (Newing et al., 2013a). In the next two chapters, this previously limited investigation is addressed directly. Temporal revenue data is analysed and how varied revenue is over time and space is demonstrated. Exploring the novel opportunities in data, relationships between supply and demand are examined providing clear insight into the impact of time and providing data capable of rectifying the inadequacies of traditional SIMs.

Chapter 4 – Temporal Fluctuations in Retail Sales

4.1 Introduction

This chapter presents the first of two novel data analysis chapters, which aim to address a key research objectives of this work by deriving greater insights into consumer behaviour and temporal components of demand using the partner's data. These insights will be used to refine the SIM (chapters seven and eight) by offering deeper understandings of temporal demand fluctuations and their impacts on store level sales, presenting opportunities for novel developments in location planning in the grocery sector.

The aim of this chapter is to examine the temporal fluctuations in revenue from a supply-side perspective, identifying the degree to which revenue changes and is impacted over time and the drivers behind these changes. Understanding the temporal fluctuations of grocery sales is likely to be an important component for accurately predicting store revenue and store performance through an improved understanding of demand and remains underrepresented in the literature. Temporal dimensions impact on store revenue in a multitude of ways. Temporal variations are an important driver in consumer spending that has a significant impact upon store revenue. A number of studies now exist in population geography articulating the importance of considering variations in population distributions over time, both daily and weekly (Bell, 2015, Martin et al., 2013, Martin, 2011a). Understanding a catchment area's population profile is important for retailing; knowing the type and number of consumers in an area and how this varies over time, for example, represents important information. Time is a driving factor of local demographic change, resulting in very varied market conditions over 24hr periods, consequently impacting on store trade and revenue (Bell, 2015, Berry et al., 2016, Martin et al., 2015, Newing et al., 2013a, Schwanen, 2004). It is therefore important to consider time when modelling consumer behaviour.

As noted above, there has been some research into temporal fluctuations of population distributions. A detailed review on spatial-temporal population modelling can be seen in Martin et al. (2015). They review previous work that has attempted to estimate 'night-time' populations represented by census data and other sources: see also Coombes (2010), (Martin et al., 2013, McPherson et al., 2006). For the purpose of this research, a daytime population represents an average day of the week, assuming that individuals are in the usual location that they typically reside in over the course of a single day, e.g. work, school or home. The actual nature of population distributions is far more fluid and complex which others have attempted to model using more discrete time scales as well as a more explicit representation of the population; for example, looking at the modelling of temporal populations within a university campus and at airports and cruise terminals (Charles-Edwards and Bell, 2013, Jochem et al., 2013). For airports, for example, these studies take into account flight times, transport statistics, arrival times at different times of the day or day of the week. For universities, term dates have been used to build up a more sophisticated representation of demographic activity at these sites. However the limitation of increased detail is in the flexibility of these models and their wider application. While more sophisticated, they have limited representation and are often being applied to single sites only. There are fewer applications implemented for a larger geographical scale. Martin et al. (2015) note that while the principal division of daytime and night-time is appropriate, particularly in terms of flexibility in reproducibility, they suggest the need for caution in over simplification of demographic groups used within these temporal periods. While a discrete representation of time limits the flexibility in representing population distributions (Martin et al., 2015), it is important to account for multiple behaviours over time to ensure a realistic framework is achieved.

The importance of time in terms of demographics is the distribution of the population and in relation to retail sales this is paramount. The distribution and actions of consumer demand and how it varies have a major impact on store performance and revenue (East et al., 1994). The ability to model the distribution of consumers accurately through time is therefore of critical importance for businesses, particularly when making location based decisions on store performance and store planning. Identifying varying sources of expenditure and the amount of available expenditure (and likewise the behaviour of the consumers both spatially and temporally) is necessary to understand when building models of retail markets. As discussed in the previous chapter, consumer behaviour varies temporally and spatially. The distribution of consumers (excluding E-commerce and M-commerce spending) is typically representative of the physical location of available expenditure at a particular site, at that point in time. Therefore by better understanding the spatiotemporal behaviour of demand types, it is possible to improve SIMs, thereby improving their predictive capacities.

From a retailer's perspective, time of day is also important at an operational level. In off-peak hours stores may operate price differentials to boost demand or use quieter periods to restock products, undertake training or perform maintenance (East et al., 1994, Gauri et al., 2008). Similarly, it has been suggested that in busy periods (at certain times of day or over certain periods in a year) stores may be able to boost staff levels or stock particular products aimed at specific customers and shopping missions (Berry et al., 2016, East et al., 1994, Newing et al., 2013a). The value and volume of products changes both spatially and temporally and can occur over a small time scale (such as hourly throughout a day) but also across a longer time scale, such as over the holidays. In the former situation, workers may buy lunch during the day or purchase food for dinner and alcohol for home at the end of the day (Berry et al., 2016). In the latter situation, tourists on holiday can boost demand in an area over a period of months as well as demonstrating very different consumption patterns compared to residents during their holidays (Newing et al., 2013a). It is important to account for temporal differences to ensure stores are adequately supplied and capable of catering to consumer demand. These operational decisions may manifest themselves through stocking more of a specific product at a certain time of day or operating larger stores due to seasonal tourism or promotional changes reacting to changes in the weather (Berry et al., 2016, East et al., 1994, Newing et al., 2013a). It is clear that retailers are sensitive to the impact of changing temporal dimensions in market conditions, which in turn impact store revenue. What is less clear is the extent to which revenue varies over time. The remainder of the chapter will present evidence of revenue change over time, providing insight into the grocery sector and aiding in improving our understanding of consumers through the analysis of temporal supply-side data.

4.2 Introduction to The partner's store data (supply-side)

In section 4.1 it was argued that time is an important factor impacting revenue. Revenue is highly variable over time and this has previously been shown over an extended time scale (Newing et al., 2013a). However, fewer studies have shown that significant variation can be observed over much shorter timescales. The rest of this chapter explores temporal sales data provided by the data partner. The temporal dataset consists of 136 individual grocery stores located within The Yorkshire and Humberside (Y+H) regions of the United Kingdom. More specifically, the dataset represents 36 supermarkets and 100 convenience stores. Fig 4.1 below outlines the spatial distribution of the store network in the region. The revenue data provided comes from a consecutive seven-day period for a full week of sales from October 2014. The dataset underwent data cleaning and manipulation from its raw state. Transactions (defined below) were grouped into ten minute intervals throughout each day to reduce noise in the data. Table 4.1 (below) shows a summary of the records contained within the dataset. The data used originates from the electronic point of sale system within each store. Every item that is scanned at the till is recorded (a transaction), along with its weight (if relevant) and cost. This is what each customer sees on their receipt when they have paid. The data are transferred to the data centre, where processing is conducted before the data are exposed to the in-house analytical community via the 'Enterprise Data Warehouse'. In the case of the data used here, the transaction summary is used containing only the key fields for the project; as seen in Table 4.1. Yorkshire and Humberside contains approximately 10% of the partner's convenience store square footage, and 5% of supermarket square footage. The convenience stores account for almost 11% of overall convenience store revenue, and supermarkets nearly 6% of supermarket revenue nationally. Given the scope of this thesis and level of data necessary to develop a detailed SIM, a sample of this Yorkshire and Humberside data (47 stores within and on the border of West Yorkshire only) will be used to build the SIM. Unless otherwise specified below the discussion and analysis presented in the next section of this thesis (section 4.3), will be derived from the temporal revenue data of these 47 stores (contained within the highlighted region), presenting insight on temporal fluctuations of revenue within West Yorkshire.

Figure 4.1 – Map showing network of all The partner's stores in Y+H represented in the temporal sales dataset, West Yorkshire sample highlighted in the circle.

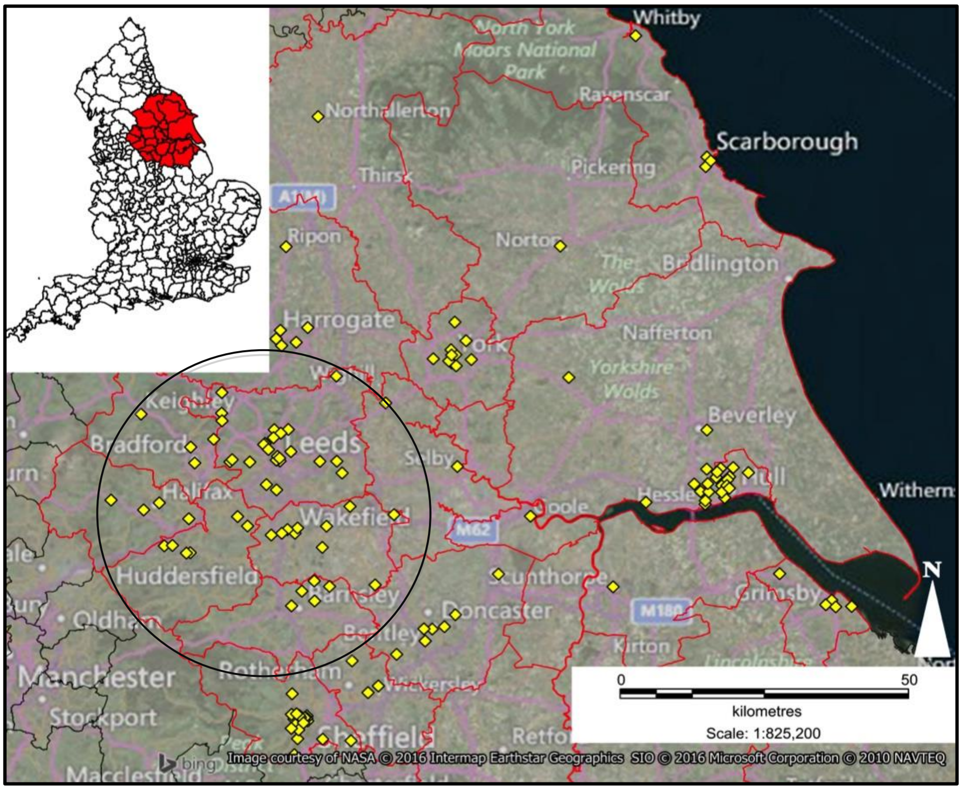


Table 4.1 - Summary of temporal sales dataset records.

Database name	**Trade profiles**	**136 stores**
Tables contains 24hr transaction records for stores in West Yorkshire between the 12th- 18th of October 2014		
Variable name	**Description**	
TRANSACTION_DATE	Data of transaction	
TRANSACTION_TIME	Time of transaction (24hr clock)	
TRANSACTION_ITEM_COUNT	Number of items purchased	
TRANSACTION_VALUE	Total value of transaction (£)	
TRANSACTION_DISCOUNT	Discount value, if any, on the transaction (£)	
SiteKey_As_JSBranch	Branch ID number. Can be used to link purchases to stores spatially.	

For the analysis, stores are allocated into two format types: **con**venience stores and supermarkets. Convenience stores are defined as having a floorspace <3000 sq./ft. and supermarkets >3000 sq./ft. (Hood et al., 2015). The partner's location planning team have classified stores on an individual basis via several 'location type' classifications within each store format type. Within West Yorkshire the following location types are used by them to classify stores:

> Supermarket location types – *Suburban High Street, Small Town, Retail Park, Out of Town Mall, Major Town Centre, Edge of Town Centre, District Centre.*
> Convenience Store location types – *Standalone, Major City Centre – Shopping Area, Major City Centre – Office District, Local Pitch, High Street.*

Analysis in the next section will be initially undertaken on the dataset by location type and store format presented above. Although location names are somewhat self-explanatory of the location type, definitions for each specific location type were not included in the data provided by the commercial partner. These location types were developed internally and are used to provide context of a store location and are well used by the partner. Furthermore, using the partner's own location type classifications means that any demand and supply side insights can be provided within an established operational system, meaning that new insights can be implemented directly into existing operations.

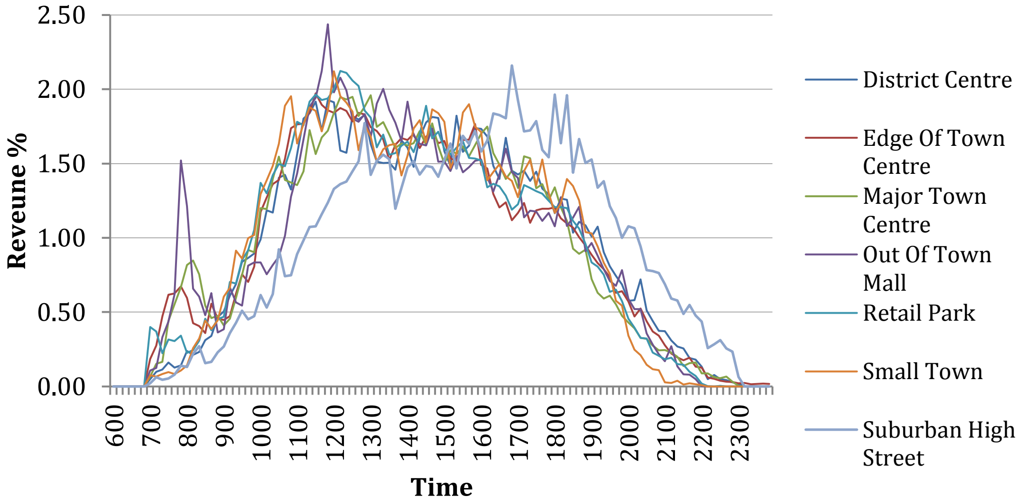
4.3 Temporal revenue analysis

In this section an analysis and discussion of the temporal revenue data is presented. Identifying any trends and patterns in the data in order to provide insights into the impact of variations in transaction value and volumes and how this varies by time of day.

Figure 4.2 - Shows revenue (as a % of total revenue) for grocery stores, shown by format type, and by hour of the day, averaged across all days in the week (6am-Midnight)

Figure 4.2 shows the percentage of total revenue made hour by hour for all stores along with format type (averaged over the week). Revenue is clearly seen, for both supermarkets and convenience stores, to vary throughout the day, as well as varying by store format type. Both formats demonstrate a different peak-time in revenue, with supermarkets peaking earlier in the day from around 11am to 1pm and then tailing off after this peak until 6pm when revenue drops more rapidly. In contrast, the major peak in revenue for convenience stores begins after 5pm lasting until around 7pm and then subsequently declining. Less pronounced peaks in revenue are seen at 9am and over the lunch period, which could be the result of consumers buying food for immediate consumption, such as breakfast and lunch respectively. For convenience stores, there is a peak in the early evening as supermarket sales begin to decline. Perhaps more people are shopping at convenience stores for their evening meal, especially on Sundays when the Sunday Trading Law limits the opening times of stores over 3000 Sqft (limiting opening trade to a total of six hours). The full extent of Sunday trade and the resulting revenue profiles are presented below. The purchasing of food at different times reflects long term transformations in consumer behaviour, such as changes in lifestyles and consumption patterns (see early discussion in Chapter two) with consumers seeking more fresh produce and undertaking smaller but more frequent shops, therefore reducing the demand for the more traditional weekly shop in a large supermarket. Similarly, this may be due to short-term changes influenced by more immediate factors such as the location of stores and consumers, the time of day and convenience of shopping with stores strategically located and planned to cater for temporal consumer needs (Baron et al., 2001, De Kervenoael et al., 2006, Hallsworth et al., 2010, Hood et al., 2015).

Figure 4.3 - Shows revenue (as a % of total revenue) for supermarket stores, disaggregated by location type, time of the day, averaged across all days in the week (6am-Midnight).



In Figure 4.3 the majority of supermarket location types mirrors the supermarket format temporal profile pattern for weekly revenue. All of the supermarket locations appear to have a peak in revenue during the day and 55% of total revenue is taken between the times of 11am and 4pm. The ‘Edge of town centres’, ‘Major town centres’ and ‘Out of town malls’ all exhibit a peak in revenue early in the morning. This is also observed to a lesser extent at the ‘Retail park’ locations, although this appears to occur marginally earlier. It is likely that the uplift in revenue during this morning period is work related and the result of consumers arriving for work at, or in close proximity to, these locations or passing through and picking up breakfast or lunch. Aside from not demonstrating a morning peak in sales, ‘Small town’ and ‘District centre’ location types follow a similar temporal profile to the other locations. Generally speaking revenue profiles, especially during the middle of the day, are comparatively similar for all supermarket types. This relationship suggests that it is likely that the observed temporal profiles are likely to be representative of most supermarket trade, i.e. the times that supermarkets receive revenue will be comparable, following similar temporal patterns over the course of a day (see Table 4.2 below as well as section 4.4 for additional analysis). The exception to this is the ‘Suburban high street’ location, which demonstrates a delayed trade profile with increasing sales occurring later in the day between 1pm and 7pm. Although sales, as with the other location types, begin to decline from 7pm onwards, the proportion of sales in this particular location type remain higher for longer, proportionally generating more revenue than the other location types up until this time. The trade profile of the ‘Suburban high street’ location is likely due to travel patterns of the surrounding catchments in this location type. The name ‘Suburban high street’ suggests a smaller shopping area with mixed residential and commercial properties (a site analysis of the data partner’s ‘Suburban high street’ location type supports this assumption) and it is likely that the extended peak of revenue at the end of the day is due to some residents leaving and returning to the area and potentially shopping after work. A previous study found that the grocery consumers who undertook a food shop ‘on the work route’, tended to favour the evening period when grocery shopping (East et al., 1994). It is reasonable to assume that consumers who continue to demonstrate more traditional shopping habits (undertaking a big weekly shop) still follow a similar decision making process. It is also plausible to suggest that some proportion of the revenue resulting in the day time peaks demonstrated in Figure 4.3 is also caused by consumers on their days off (again, changing lifestyles and irregular working hours may be an influential factor here). There has been an increasing number of families with single income earners in addition to general wages increasing, it may also reflect those who are unemployed or older, retired consumers (see previous chapter) (Hallsworth et al., 2010). There have been dramatic changes to working hours in addition to evolving working environments from more conventional operations in recent years, working in multiple locations (such as the home) and over varied hours (Admin, 2016).

Table 4.2 - Correlation matrix of Supermarket location types revenue profiles (SPSS output).

		District_ Centres	Edge_of_ Town	Major_Town_ Centre	Out_of_Town	Retail_Park	Small_Town	Suburban_ High_Street	Average
District_Centres	Pearson Correlation	1	.967**	.958**	.925**	.973**	.971**	.855**	.989**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000
	N	108	108	108	108	108	108	108	108
Edge_of_Town	Pearson Correlation	.967**	1	.972**	.953**	.985**	.967**	.767**	.983**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000
	N	108	108	108	108	108	108	108	108
Major_Town_ Centre	Pearson Correlation	.958**	.972**	1	.948**	.970**	.959**	.795**	.982**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000
	N	108	108	108	108	108	108	108	108
Out_of_Town	Pearson Correlation	.925**	.953**	.948**	1	.936**	.908**	.768**	.958**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000
	N	108	108	108	108	108	108	108	108
Retail_Park	Pearson Correlation	.973**	.985**	.970**	.936**	1	.980**	.770**	.984**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000
	N	108	108	108	108	108	108	108	108
Small_Town	Pearson Correlation	.971**	.967**	.959**	.908**	.980**	1	.790**	.979**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000
	N	108	108	108	108	108	108	108	108
Suburban_High _Street	Pearson Correlation	.855**	.767**	.795**	.768**	.770**	.790**	1	.852**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000
	N	108	108	108	108	108	108	108	108
Average	Pearson Correlation	.989**	.983**	.982**	.958**	.984**	.979**	.852**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	
	N	108	108	108	108	108	108	108	108

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.4 - Shows revenue (as a % of total revenue) for convenience stores, disaggregated by location type, time of the day, averaged across all days in the week (6am-Midnight).

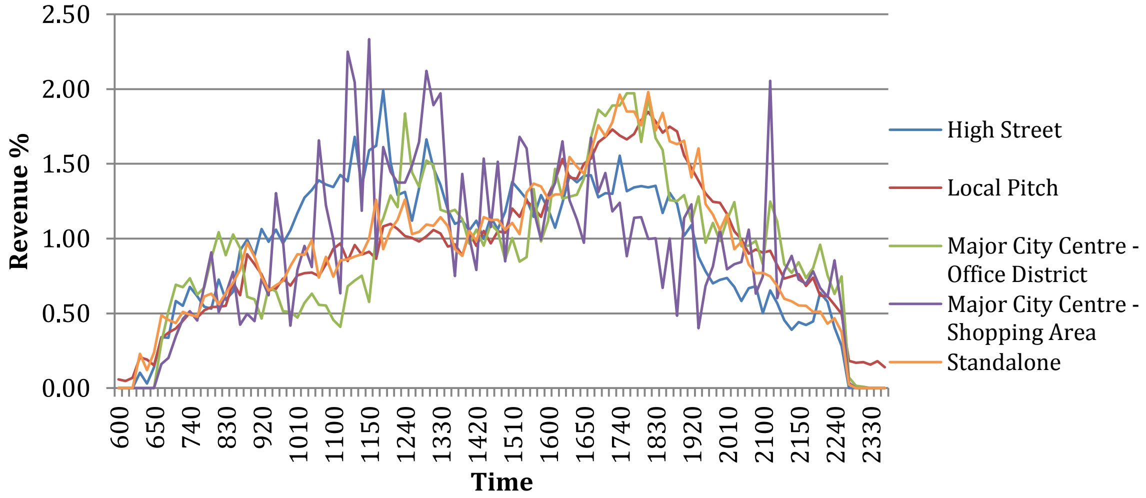


Figure 4.4 shows the trade profile for all convenience store locations in the study area. Unlike the supermarket location types, trade appears to be less consistent (hour by hour) across the convenience location types with more definable periods of revenue change. This implies that convenience stores are more likely to have varying periods of busy trade and quiet trade compared to a more consistent flow of revenue observed at the supermarket location types (see Figure 4.3). This is shown via the inconsistent revenue profiles. The data shows that there is a higher tendency for convenience store location types to have dissimilar trade profiles and that

the more profitable periods are split up, occurring at different times in the day. However, some patterns can be observed between individual location types, suggesting the influence of similar temporal and consumer components. Of all the location types, 'Standalone' and 'Local pitch' location types are the most alike with an r-value of 0.97, demonstrating a very consistent trade profile over time, and are most profitable in the early evening. This may be an indicator of consumer behaviour (shopping at the end of the working day) or similarly increasing demand (for example, residents returning home). The location type 'Major city centre – office district' appears to have the most distinctive temporal periods. Revenue peaks at the following times: 8-9:30am, 12:30-1:30pm and 5-6:30pm. It is likely that these times are linked to key commuting periods to and from work (as well as the lunch time break) with customers buying breakfast or food for lunch and then evening meals at the end of the day - a behaviour noted previously (see Chapter Two). It is clear that this location type's trade profile (and thus revenue) is greatly influenced by the presence of a day time work force, an assumption that is reflected on in Figures 4.7, 4.8, 4.9 and 4.10, which show the existence of morning and lunch time trade on weekdays that are less pronounced on weekends. It is more than likely that workers are the main source of revenue for stores in this location (a spatial analysis found that within 500 metres of a 'Major city centre – office district' store, the workplace (WP) population was approximately 38,000) and the trade profile is indicative of this. In contrast, the remaining two location types, 'High street' and 'Major city centre – shopping district', appear to take their highest portion of revenue leading up to lunch time (both taking 26% of their total revenue during the peak period), and experience a smaller uplift in sales in the evening with declining revenue seemingly coinciding with non-food store trading hours. This suggests that the consumers in these locations follow different temporal patterns, behaving and consuming differently to residential and workplace consumers. This has also been suggested by Roy (2004) and Schwanen (2004), implying that it may be necessary to factor non-residential and non-worker consumer demand into any future model.

Figure 4.5 - Shows percentage of cumulative revenue (£) for supermarkets, disaggregated by location type, demonstrating the progression of sales by hour of the day.

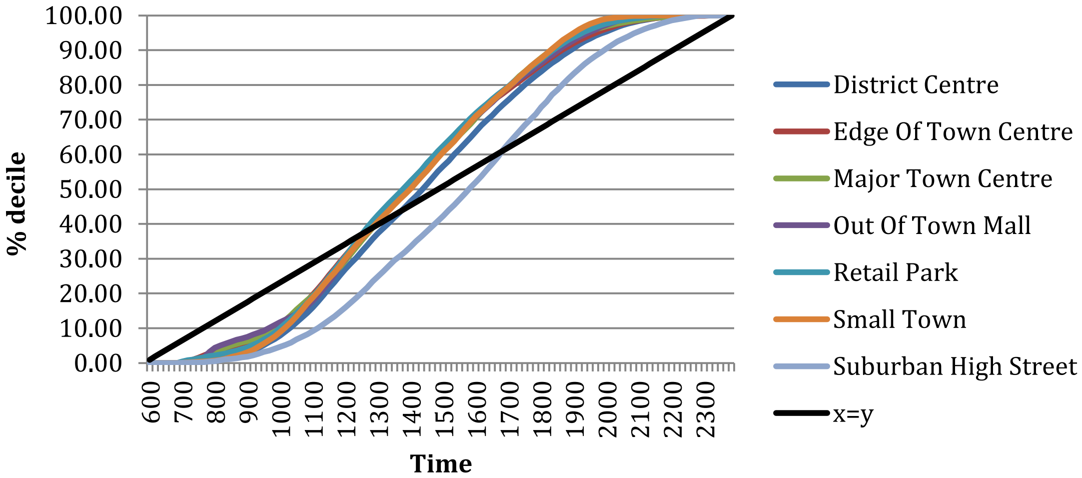
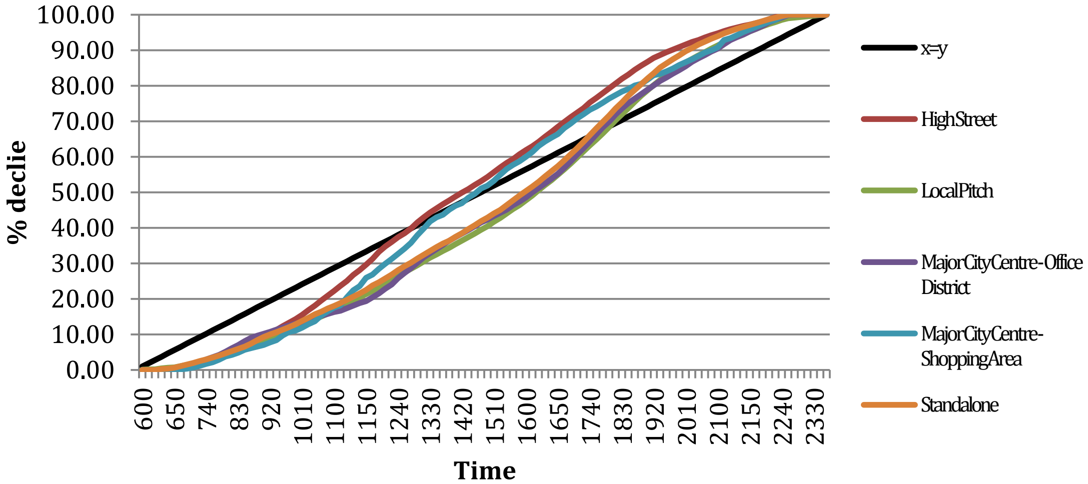


Figure 4.6 - Shows percentage of cumulative revenue (£) for convenience stores, disaggregated by location type, demonstrating the progression of sales by hour of the day.



Figures 4.6 and 4.7 present the cumulative revenue profiles of both the supermarket and convenience store location types (respectively), as a percentage over time. In both figures the x=y trend line represents a perfect positive relationship between revenue on the Y-axis and time on the X-axis, i.e. revenue increases at a constant rate all day. The data presented in these figures further highlights the differences between supermarkets and convenience store locations, with supermarkets accumulating a higher proportion of their revenue earlier in the day. It is also possible to see that 'Suburban high street' locations (as previously observed in Figure 4.3) generate more revenue later in the day with a delay of around two hours compared to the other supermarket locations, e.g. reaching 10% of revenue at 11am compared to other supermarkets

reaching 10% of revenue at or before 10am and 50% at 15:30 compared to 13:30. Similarly, ‘High street’ and ‘Major city centre – shopping district’ convenience store locations (as seen in Figure 4.4) generate the majority of their revenue early in the day (accumulating 50% by 14:30) compared to the other convenience location types which reach 50% of revenue at 16:00. Supermarkets demonstrate similar cumulative revenue profiles (Figure 4.5), again suggesting that shopping behaviour is comparatively similar across supermarket locations. On the other hand, convenience store locations (Figure 4.6) are more varied with observed periods of rapid revenue gains as well as slower periods, which may reflect the transient nature of convenience stores demand and the level of temporal fluctuation that occurs.

Stores classified as the supermarket format predominantly generate the partner’s revenue; supermarkets take over 84% of total grocery revenue for the grocery retailer in West Yorkshire. The biggest revenue source by supermarket location type is ‘Edge of town centre’ locations, generating over 27% of total supermarket income, followed by ‘retail park’ locations. The biggest revenue source by convenience store location type is the ‘local pitch’ followed by the ‘standalone’ location type, respectively taking a share of 57% and 23% of total grocery revenue from the convenience store format. These values are not unexpected as these location types contribute the highest proportions of floorspace. It is important to consider how different the temporal revenue profiles (shown in Figures 4.2, 4.3, and 4.4) would look if raw revenue values (£) were used instead of revenue as a percentage, being dominated by the higher earning location types.

Table 4.3 (below) provides a breakdown of grocery revenue by day of the week for all stores and by format and location type. The biggest share of revenue is made on Fridays and Saturdays (18.2% and 19.1% respectively), beginning to increase as the end of the week approaches (Thursdays) and then dropping notably on Sunday. More specifically, Saturdays are the biggest revenue earner for supermarkets in contrast to Fridays for convenience stores. The Friday and Saturday trend is more prevalent in supermarkets, which are considerably lower during the week and on a Sunday. The daily distribution in the share of revenue when broken down by supermarket locations again, proportionally, follows the same pattern as the overall results with Saturdays observed to be the main day of the week for grocery shopping. East et al. (1994) provided the opportunity to compare results and demonstrate evolving grocery shopping behaviour, presenting comparable data on daily shopping revenue shares (%) attained from a consumer research survey conducted over a four week period. In their study they outline a similar distribution of grocery shopping throughout the week, observing that consumers preferred to shop at the end of the week, favouring Friday, Saturday and to a lesser extent Thursday. The major difference between daily shopping habits then and now is Sunday shopping. In their study they observed a share of only 3% of total weekly revenue on Sunday compared to a total of 10.8% observed in the partner’s data. This difference is likely in part due to changes in lifestyle and to the more liberal Sunday trading laws introduced in the early 1990s (De Kervenoael et al., 2006). This theory is supported by the evidence collected in the ONS Time Use Survey conducted in 2000, after the changes in trading laws, which also demonstrated an increase in grocery shopping on a Sunday (Ipsos-RSL, 2003). Contrastingly, while Friday provides the biggest share of daily revenue for convenience stores, the level of variance throughout the week is very small (0.05) and the increased Friday share is limited to a much smaller percentage change. This suggests that consumer behaviour and demand remain relatively consistent throughout the week for convenience stores with a standard deviation of 0.22 (and proportionally this appears to be the case across all location types), demonstrating consistent sales regardless of day.

Table 4.3 - Daily share of the total weekly revenue (%) by format and location type.

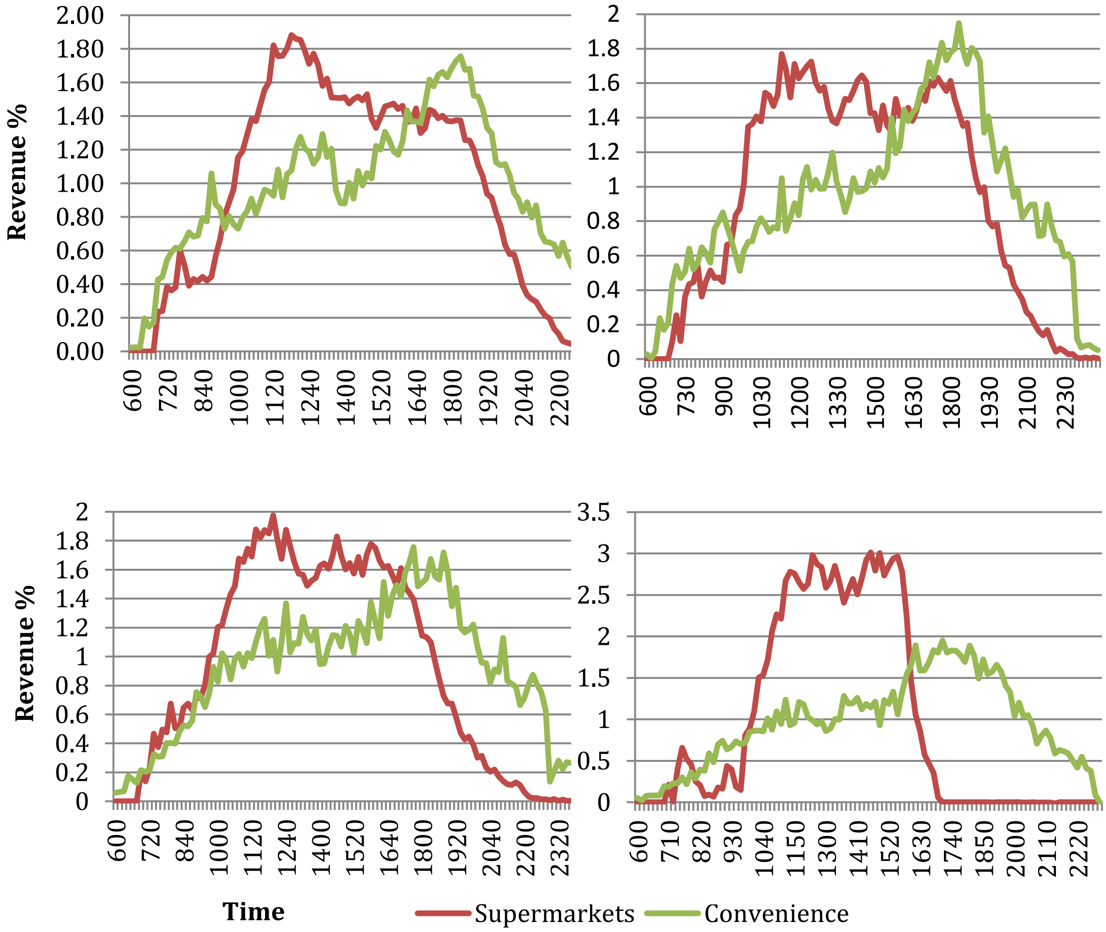
Store category	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Grand Total
All Stores	12.7	12.4	12.4	14.4	**18.2**	**19.1**	10.8	100.0
Supermarkets	10.7	10.2	10.3	12.2	15.5	**16.6**	8.5	84.0
Convenience	2.1	2.2	2.1	2.2	**2.7**	2.5	2.3	16.0
District Centre	1.37	1.31	1.28	1.52	1.86	**2.15**	1.16	10.7
Edge Of Town Centre	2.85	2.71	2.70	3.48	4.46	**4.72**	2.43	23.4
Major Town Centre	1.83	1.82	1.84	2.06	2.56	**2.58**	1.18	13.9
Out Of Town Mall	0.98	0.87	0.92	1.06	1.28	**1.43**	0.87	7.4
Retail Park	2.46	2.33	2.42	2.88	3.76	**4.09**	2.03	20.0
Small Town	0.81	0.79	0.78	0.91	1.19	**1.22**	0.65	6.4
Suburban High Street	0.39	0.33	0.32	0.32	0.39	**0.40**	0.21	2.4
High Street	0.21	0.22	0.22	0.22	**0.29**	0.26	0.18	1.6
Local Pitch	1.18	1.26	1.21	1.25	**1.56**	1.43	1.38	9.3
Major City Centre - Office District	0.08	0.09	0.09	0.10	**0.13**	0.11	0.07	0.7
Major City Centre - Shopping Area	0.11	0.10	0.09	0.10	**0.11**	0.09	0.07	0.7
Standalone	0.47	0.52	0.50	0.51	**0.65**	0.59	0.56	3.8

While Table 4.3 provides useful insight and detail on daily revenue shares in grocery store locations, it is not a suitable indicator of daily performance of store type and location type, due to the wide degree of variation in total revenue being made resulting from the differences in store size and numbers of stores for location types. As shown in Table 4.4, which shows daily store performance in £/Sqft for all store and location types, certain location types which were previously observed to contribute a relatively small share in daily revenue percentages are proportionally generating higher levels of revenue when we take into account total floorspace as well as total daily revenue amounts. For instance, previously in Table 4.3, the convenience store location type 'major city centre – office district' was shown to contribute the smallest proportion of revenue across the partner's store network within West Yorkshire. However, as shown in Table 4.4 this location type is by far the best performing store (based on trade intensity - TI) on each day of the week making the highest volume of revenue per square foot of floorspace. This suggests that, *ceteris paribus*, the major city centre – office district location type, and in general all convenience store location types, have the highest volumes of trade and provide the partner with the greatest level of return per square foot of floorspace for each day and overall (across the entire week), generally outperforming supermarkets TIs. This pattern is particularly seen in 'major city centre – shopping area' stores and likewise also observed in 'suburban high street' supermarket locations, which represented a very low portion of total weekly revenue (2.4%). The additional insight Table 4.4 offers is the potential to comparatively analyse daily sales (based on a stores' TI performance) across all store types as well as exhibiting that total revenue volumes are not always the most appropriate measure of store performance.

Table 4.4 - Daily store performance in £/Sqft (trade intensity - TI) for all format and location types.

Store category	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	Weekly
All Stores	2.91	2.82	2.83	3.29	4.17	4.36	2.46	22.83
Supermarket	2.73	2.60	2.63	3.13	3.97	4.24	2.18	21.49
Convenience	4.37	4.65	4.50	4.62	5.81	5.28	4.78	34.00
District Centre	3.45	3.29	3.20	3.81	4.67	5.41	2.92	26.76
Edge Of Town Centre	2.73	2.61	2.60	3.35	4.29	4.54	2.33	22.45
Major Town Centre	2.49	2.48	2.50	2.81	3.49	3.51	1.61	18.88
Out Of Town Mall	3.07	2.73	2.86	3.32	3.98	4.45	2.72	23.14
Retail Park	2.43	2.30	2.40	2.85	3.73	4.05	2.01	19.78
Small Town	2.44	2.39	2.36	2.75	3.59	3.68	1.97	19.18
Suburban High Street	**5.24**	**4.53**	**4.32**	**4.34**	**5.31**	**5.37**	**2.91**	**32.03**
High Street	3.03	3.22	3.22	3.20	4.12	3.75	2.52	23.06
Local Pitch	4.65	4.95	4.77	4.91	6.16	5.63	5.42	36.49
Major City Centre - Office District	**6.08**	**6.68**	**6.41**	**7.08**	**9.02**	**8.10**	**4.94**	**48.30**
Major City Centre - Shopping Area	6.05	5.44	5.20	5.75	6.10	5.07	3.93	37.55
Standalone	4.09	4.49	4.33	4.38	5.64	5.11	4.86	32.89

Figure 4.7 - Shows revenue (£) as a percentage of total revenue by time of day, for grocery stores, by format type, based on revenue averaged over (A – *top left*) Monday through to Thursday, (B – *top right*) Friday, (C – *bottom left*) Saturday, (D – *bottom right*) Sunday.



Based on the data presented in Table 4.3 and 4.4 revenue profiles were grouped into four diurnal periods with distinguishable revenue totals and is as well as socio-economic perceptions, i.e. days off, end of the week (East et al., 1994). These are: Monday to Thursday, Friday, Saturday and Sunday (Figure 4.7), which are presented as hourly revenue profiles for supermarket and convenience store formats (expressed as a percentage of total format revenue in that period). Individual location types for both supermarkets and convenience stores were also reviewed separately for each daily period. However, the additional disaggregation in store data failed to provide any additional insight or detail of revenue profiles by hour of the day. Instead, the decision was made to focus analysis on average hourly revenue variation over days of the week for each store format type. Aside from the obvious differences between supermarkets and convenience stores, which have previously been discussed above, revenue profiles for both store format types, (in particular convenience stores) are relatively consistent throughout the week. They demonstrate a strong correlation between daily revenue profiles over time for each daily period, excluding the difference in the hourly trade profiles for supermarkets on a Sunday (caused by restricted opening hours): see Tables 4.5 and 4.6. The early peak in revenue at 7:40 am on Sunday, before supermarkets open, is likely to be the result of operational processes, such as transactions carried over from the previous day, internal charges or online sales (as indicated by staff from the commercial partner).

However, there are some less obvious variations in the hourly profile patterns for each format type that occur throughout the week. These include a more distinct morning and lunchtime peak in convenience stores Monday through to Friday. On Saturdays and Sundays there is less evidence of a morning peak in the convenience stores and while there is a slight lunchtime increase, it is less than those exhibited on a weekday. This difference is likely due to the presence of workplace demand, which Berry et al. (2016) likewise suggest is a factor influencing weekday/weekend convenience store revenue profiles. Supermarket trade is very similar on a Friday and Saturday exhibited by Pearson's product moment correlation coefficient with an r value of 0.958: this similarity is not unexpected as they represent the main shopping days for consumers. Consumer behaviour is likely to be very similar on those days, with revenue almost consistently similar throughout the main part of the day (10am-6pm). A reduced tendency to do a weekly shop at the start of the week (East et al., 1994) is reflected in the lower sales Monday-Thursday (Table 4.3). However, similar to the convenience stores, there is evidence of a lunchtime peak (potentially caused by workers) although revenue tails off after this period. Overall, trade profiles are relatively consistent and the only apparent distinction is the revenue totals in pounds for each day being spent (see Table 4.3). This suggests that while consumers tend to spend more or less depending on what day of the week it is, or the format type of the store they are shopping in (i.e. the total value spent in store), the temporal pattern (i.e. the time that consumers shop and grocery retailers make their money throughout the day) varies less. It remains relatively consistent across different diurnal periods, shown via the strong correlations in Table 4.5 (excluding supermarkets on Sundays) and 4.6. This finding suggests that by accounting for the average temporal fluctuations that occur over time (assuming realistic weekly grocery spending estimates), models can capture the important and cyclical elements of consumer behaviour and demand that occur on a more discrete daily basis.

Table 4.5 - Correlation matrix of daily revenue profiles in supermarkets

		Monday to Thursday	Friday	Saturday	Sunday
Monday to Thursday	Pearson Correlation	1	.986**	.966**	.753**
	Sig. (2-tailed)		.000	.000	.000
	N	108	108	108	108
Friday	Pearson Correlation	.986**	1	.959**	.691**
	Sig. (2-tailed)	.000		.000	.000
	N	108	108	108	108
Saturday	Pearson Correlation	.966**	.959**	1	.807**
	Sig. (2-tailed)	.000	.000		.000
	N	108	108	108	108
Sunday	Pearson Correlation	.753**	.691**	.807**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	108	108	108	108

**. Correlation is significant at the 0.01 level (2-tailed).

Table 4.6 - Correlation matrix of daily revenue profiles in convenience stores

		Monday to Thursday	Friday	Saturday	Sunday
Monday to Thursday	Pearson Correlation	1	.962**	.924**	.940**
	Sig. (2-tailed)		.000	.000	.000
	N	108	108	108	108
Friday	Pearson Correlation	.962**	1	.918**	.940**
	Sig. (2-tailed)	.000		.000	.000
	N	108	108	108	108
Saturday	Pearson Correlation	.924**	.918**	1	.938**
	Sig. (2-tailed)	.000	.000		.000
	N	108	108	108	108
Sunday	Pearson Correlation	.940**	.940**	.938**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	108	108	108	108

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 4.8 - Average basket value (£) for all format and location types.

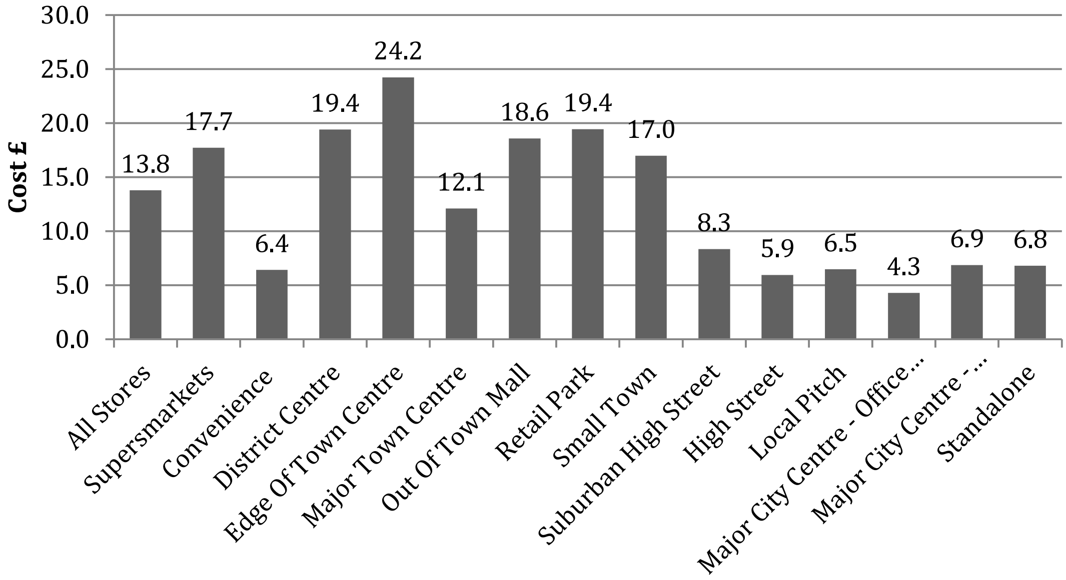


Figure 4.9 - Average basket size (No. of items) for all format and location types.

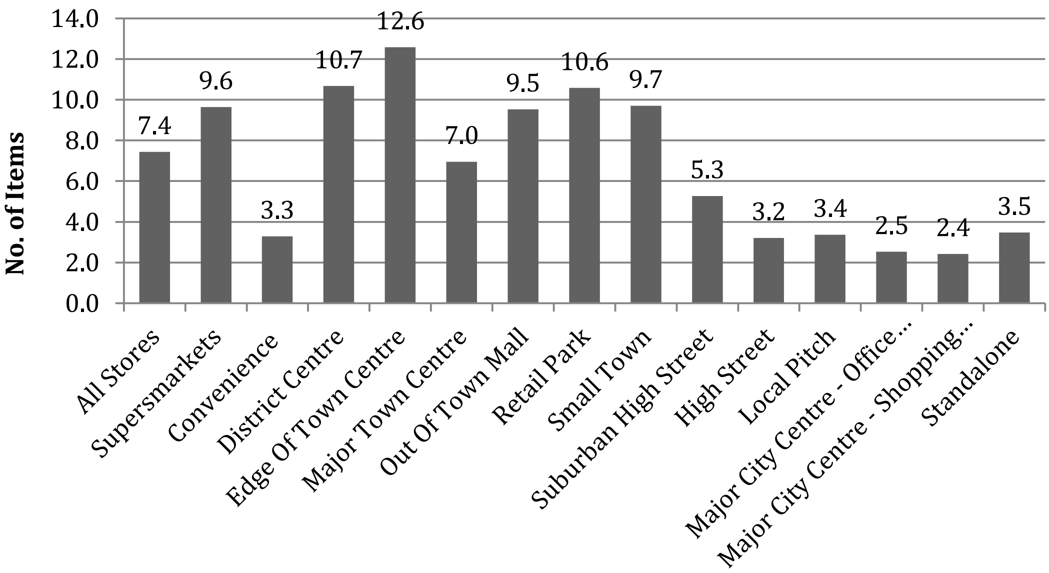


Figure 4.8 shows the average 'basket' (transaction) value (averaged over a week's worth of sales) in each format and location type. Figure 4.9 shows the average number of items in each basket over the same period, again, for all format and location types. It appears that the average basket size is always lower in convenience stores than in supermarket stores, and that the average basket size and value is almost three times higher in supermarkets. As shown previously, in chapter two, the socioeconomic profile and lifestyles of consumers in recent years have changed, with fewer people deciding to shop weekly. Convenience stores have been the

grocery industry's response to this shift in demand, offering a smaller selection of fresh produce to cater to this changing customer need (for smaller and more frequent purchases – food for now, food for later or top up shops), with consumers wanting a bigger range of products and purchasing less frequently (Clarke, 2000, Dashwood, 2013a, Hallsworth et al., 2010, Hunter, 2004). Other motivations behind smaller basket sizes include a consumer's inability to plan ahead, making impulse or last minute purchases or taking advantage of price discounts (Bell and Lattin, 1998, Desai and Talukdar, 2003).

Interestingly, suburban high street location types, although classed as a supermarket, have a much smaller basket size and value (smaller in fact than the other supermarket location types and more similar to convenience store locations). This may be a result of consumers for this location type exhibiting shopping behaviour more akin with those observed in smaller convenience stores (see chapter two). The revenue profile for suburban high street locations (Figure 4.3) is somewhat similar to the average revenue profile of convenience stores (Figure 4.2), which may also indicate this behaviour. A further investigation of suburban high street locations, using loyalty card product data, found that customers in these locations were purchasing similar product categories to the products being purchased in convenience stores (cold convenience, confectionary, crisps & snacks, meal solutions and soft drinks) demonstrating a strong similarity and further supporting this observation. In other words consumers are using some larger store types as convenience stores. It also appears that a relationship between basket value and size and shopping behaviour at certain store types is spatially evident too. Stores located in city and town centres have smaller baskets, with consumers spending less and buying fewer items. In many of these locations parking is limited or not available. Therefore consumers are often on foot, using public transport and maybe combining shopping trips. It is likely that the resultant behaviour, of smaller and lower cost baskets, is partially in response to consumers being restricted by the number of items they are able to carry as well as what they may be doing next (see Chapter two).

In this section it has been shown, through the analysis of temporal data, that it is possible to identify temporal patterns in store revenue and store level performance indicators, providing greater insights into temporal grocery shopping behaviour. This has also demonstrated that it is possible to infer general observations regarding consumers in general, as well as more specific consumer types (e.g. workers) from the data. The analysis has revealed that temporal variations occur on an hourly basis throughout the day and that behaviour remains relatively consistent over different days of the week, suggesting habitual behaviour for grocery shopping. Whilst on certain days relatively higher volumes of revenue are made, there are some small differences between weekday and weekend temporal revenue profiles. The overall patterns of consumption vary less on a day-to-day basis, with the highest degree of temporal fluctuations, as shown, occurring on an hourly basis. The data presented has outlined similar temporal profiles across location types, particularly convenience store locations in major city centres, thus demonstrating very similar patterns of trade as well as similar temporal profiles across format types. As well as similar supply-side behaviour regarding trade, revenue and performance patterns, there are obvious and consistent temporal components that occur throughout the day across all locations and format types. This suggests that distinctive, reliable and repetitive behaviours are exhibited by consumers and that various and regular periods of demand occur throughout the day, which could be used to identify these fluctuating demand sources (resulting in periods of high and low trade) to develop more accurate location models later in this thesis.

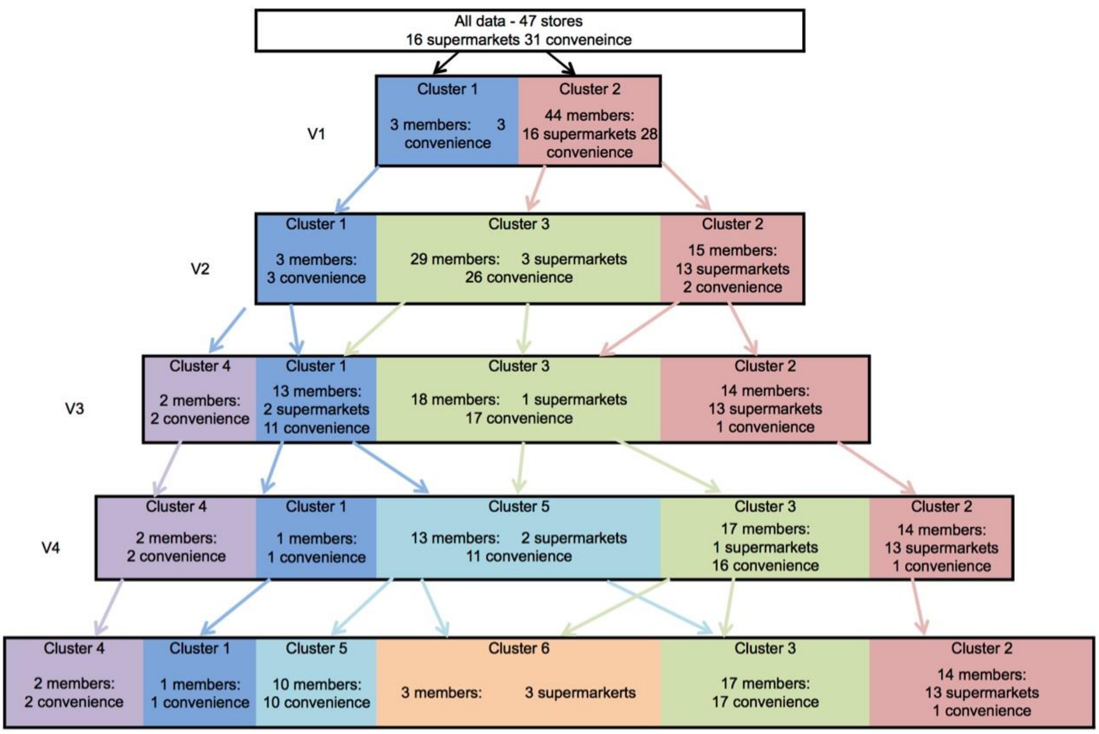
It is clear from the analysis presented above that there are obvious similarities in trade profiles and consumption behaviour between location types as well as between format types by hour of the day and to some extent by day of the week. Therefore in order to better identify the temporal patterns of trade in grocery stores and subsequently provide a clearer picture of temporal demand fluctuations, affording insight into potential sources of demand (improving model fit and performance), a cluster analysis of the data is presented next. By undertaking a cluster analysis it is possible to group similar stores, based on their temporal revenue profiles, highlighting similar temporal components in the data. The presence of any 'common' temporal patterns within clusters and across clusters will then be used to make inferences regarding the reasons behind these fluctuations in grocery demand (see Section 4.4 for details of cluster methods and the methodological practices followed to achieve this). While the data partner already has an existing segmentation of stores by store and location type, this does little to indicate the temporal profiles of stores. Therefore, this cluster analysis is necessary as it will provide novel detail regarding temporal components for stores and may also improve the distinction between the stores.

4.4 Cluster Analysis of Temporal Data

Cluster analysis is the classification of objects into categories based on the existence of similar characteristics, running iterative versions until an optimal number of clusters are found. The categorisation of objects is a useful application for numerical analysis, enabling single objects to be grouped together to identify similarities within the data (Everitt et al., 2001). In this particular application, dealing with temporal sales data in the grocery sector, it allows for further analysis building on the observed patterns and trends in the previous section. The desired outcome and rationale for the cluster analysis are to identify trends and similarities using temporal data on store level revenues by hour of the day, at an individual store level (not just by format type or location types) that prior to clustering may not have emerged. The reasoning for applying clustering for this thesis, is that it will classify stores into a smaller number of groups, aggregating stores with similar temporal revenue patterns together, making the process of identifying key temporal fluctuations, simpler. Secondly, traditional segmentation techniques, which are typically done using store missions, appeared to have little bearing on the diurnal sales profile of a store. This is being done in addition to the location type analysis in the previous section. This is because the location types developed by the commercial partner are typically used to support operational decision-making, but may not reflect the temporal difference between store revenue profiles, thus clusters may exhibit novel relationships. Thus, this approach is not only novel, but also it is likely to offer a new perspective and insight into temporal store sales that may offer insight into the drivers of sales fluctuations. The grouping of stores, which in this case provides additional analysis for understanding temporal fluctuations and the interrelationship between supply and demand, provides insight that can be used to refine location planning techniques (see chapter nine). Prior to clustering, the data was processed and consolidated into a smaller and simpler form. Using the temporal data, cumulative revenue was presented in 10-minute intervals starting from 6am (the point from which stores begin to open). The point at which cumulative revenue, in each store, exceeded each percentage decile, e.g. the time at which revenue exceeded 10%, 20% etc. for all stores and for weekdays and weekends was extracted. The extracted time periods resulted in ten data points for each store, which detailed the temporal transition of revenue and was then used for the cluster analysis.

The clustering method used was a K-means cluster (which groups objects into a predetermined number of clusters in which each object belongs to the cluster with the closest mean). It has been widely cited in the past and represents one of the most popular clustering techniques (Aggarwal, 2013, Everitt et al., 2001, Hanghang and Kang, 2013, Lingras et al., 2005). A limitation of the approach is that the K-means process does not specify the optimum number of clusters to choose for the data. There are several methods for determining the optimal number of clusters: it is possible for instance, to run a K-means scree plot in coding packages such as R, or dendograms, for example. Five versions of the K-means cluster were run using 2, 3, 4, 5 and 6 cluster groups in order to establish how robust clusters were, and to determine the best cluster solution which is shown in Figure 4.10. The optimal number of cluster groups based on temporal revenue was determined as four, which produced a k cluster solution with minimum variability within clusters and maximum variability between clusters. This was validated via an examination of the means for each cluster. ANOVA tests (which compare the means between clusters and determine if the means are significantly different from each other) were run on the temporal data used to create the clusters for each version (for example using 2-clusters, 3-clusters etc.) (Cunningham et al., 2013, Pallant, 2010). The result of the ANOVA tests showed that there was a difference between the means for each cluster version. While the ANOVA test identifies whether clusters differ, it cannot identify which clusters differ. The Tukey post-hoc test was subsequently used, which identifies which of the clustering version are significantly different and thus the best number of clusters to use (Cunningham et al., 2013, Pallant, 2010). This analysis revealed that cluster versions with 5 and 6-cluster solutions did not produce statistically significantly different clusters from each other. While the 2-cluster solution was different at all times, 4-cluster solutions were shown to represent the largest number of statistically significant different clusters (a full summary of the statistical outputs can be seen in Table 12.1 in the appendix). The outputs of the k-means clustering are presented below. Whilst there was some difference between daily revenues, particularly weekday and weekend revenue profiles, temporal revenue profiles were more similar than they were different. Standard deviation was used as a measure of fit for weekly clustering of daily profiles (Table 4.8 and Figure 4.12) suggesting that average weekly revenue was sufficiently representative of a typical day in the week's hour-by-hour sales. Consequently, K-means clustering was repeated, this time for a full week of temporal revenue data. In order to be confident in the decision to represent potential daily revenue patterns on a weekly basis, weekly cluster revenue profiles were compared to weekday and weekend temporal revenue for all stores and their relevant weekly cluster. As well as a flow diagram for weekly clustering membership solutions, presented below are the cluster membership tables (Table 4.7), mean cluster revenue profiles (Figure 4.11) and the assessment of the difference between individual store revenue (for weekly, weekday and weekend revenue) and their cluster profile using standard deviation.

Figure 4.10 - Flow diagram of cluster solutions for all K-means cluster versions of weekly temporal revenue data.



4.5 Cluster analysis outputs

In this section the outputs of the clustering are presented, followed by a discussion of the results presented in section 4.6. Here, observations between temporal fluctuations in revenue and temporal components of demand within cluster types by time of the day, location and format type, location setting and by day of the week are considered.

Table 4.7 - Cluster membership of weekly store revenue for stores including store format and by the partner's location type classifications.

ID	Location type	Format	Cluster group
2190	Major Town Centre	Supermarket	1
2222	Suburban High Street	Supermarket	1
4218	Standalone	Convenience	1
4239	Local Pitch	Convenience	1
4282	Major City Centre - Office District	Convenience	1
4480	Local Pitch	Convenience	1
4717	Local Pitch	Convenience	1
4725	Local Pitch	Convenience	1
4770	Local Pitch	Convenience	1
4774	Standalone	Convenience	1
4785	Standalone	Convenience	1
4926	Local Pitch	Convenience	1
4950	Local Pitch	Convenience	1
695	Out Of Town Mall	Supermarket	2
778	Edge Of Town Centre	Supermarket	2
793	Edge Of Town Centre	Supermarket	2
814	Retail Park	Supermarket	2
827	Major Town Centre	Supermarket	2
847	Edge Of Town Centre	Supermarket	2
867	Retail Park	Supermarket	2
897	Retail Park	Supermarket	2
2004	District Centre	Supermarket	2
2021	Small Town	Supermarket	2
2189	Small Town	Supermarket	2
2258	Major Town Centre	Supermarket	2
2295	Retail Park	Supermarket	2
4836	High Street	Convenience	2
2208	District Centre	Supermarket	3
4686	Standalone	Convenience	3
4687	Local Pitch	Convenience	3
4713	Standalone	Convenience	3
4719	High Street	Convenience	3
4730	High Street	Convenience	3
4744	Local Pitch	Convenience	3
4754	Local Pitch	Convenience	3
4757	Local Pitch	Convenience	3
4763	Local Pitch	Convenience	3
4772	Local Pitch	Convenience	3
4778	Local Pitch	Convenience	3
4779	High Street	Convenience	3
4780	Standalone	Convenience	3
4784	Local Pitch	Convenience	3
4804	Standalone	Convenience	3
4807	Standalone	Convenience	3
4813	Major City Centre - Shopping Area	Convenience	3
4458	Local Pitch	Convenience	4
4736	Local Pitch	Convenience	4

Figure 4.11 - Mean temporal revenue profile by hour of the day for each cluster group.

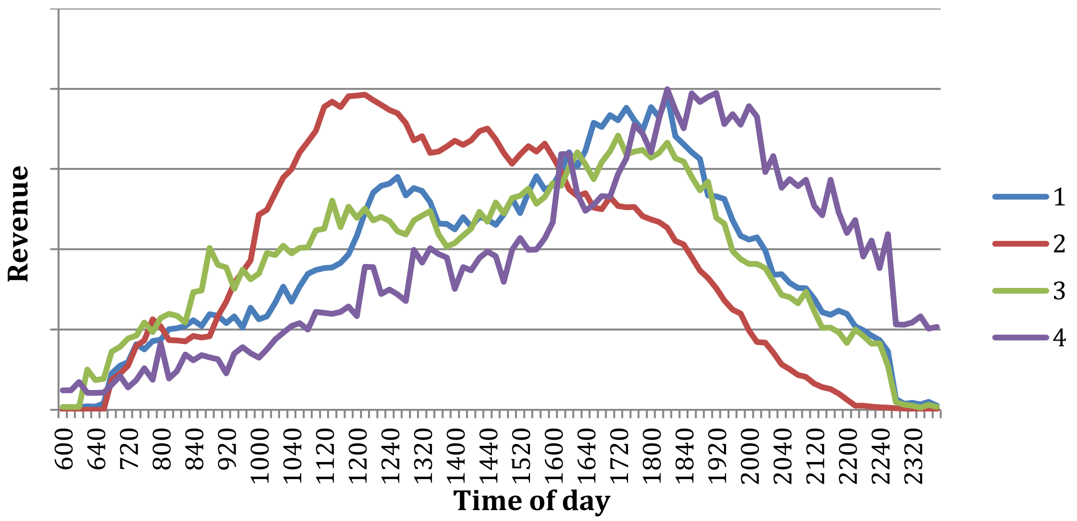
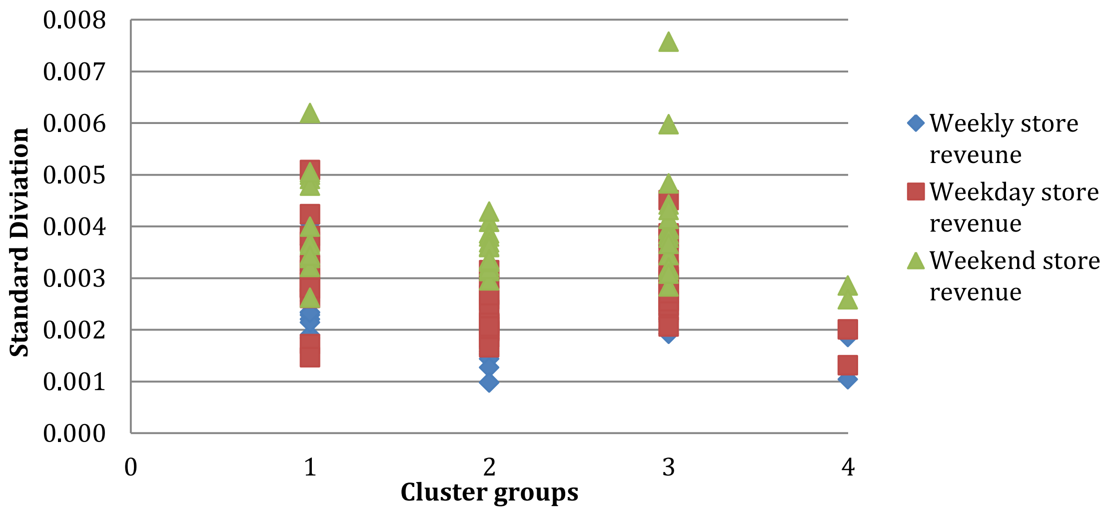


Table 4.8 - Mean standard deviation between all stores and cluster groups by weekly, weekday and weekend hour-by-hour revenue.

	1	2	3	4
Weekly revenue	0.003	0.002	0.003	0.001
Weekday revenue	0.003	0.002	0.003	0.002
Weekend revenue	0.004	0.004	0.004	0.003

Figure 4.12 - Standard deviation (STD) between weekly, weekday and weekend revenue for all stores and their cluster revenue profiles at an individual level.



The low STD values and limited spread (seen in both Table 4.8 and Figure 4.12) suggest that cluster group mean revenue profiles are indicative of each store within the cluster individual revenue profiles i.e. the mean profiles in Figure 4.11 are representative of store revenue whether it is considered separately as either weekly, weekday or weekend revenue, as very low standard deviations were found. This analysis allows for a good level of confidence in the representation of stores by these clusters, supporting the decision to use a mean weekly revenue profile. Using clusters it is potentially simpler to identify common temporal fluctuations that may provide insight into both consumer behaviour and changing demand, from a smaller sample of revenue profiles limiting the potential for human error through a reduced number of data (i.e. using four cluster profiles to represent all stores, instead of 47 individual store profiles for analysis).

4.5.1 Preliminary observations of cluster outputs

Discussion of the patterns and interpretations of the findings is presented in Section 4.6. However, first loose cluster descriptions are presented based on the observed temporal profile patterns presented above (Figure 4.11), together with known features about member stores. The potential use of cluster groups as a tool for predicting temporal trade patterns (i.e. the time of the day when stores tend to make money, but not the amount of money) will be considered in Chapter nine.

- Cluster 1: *Workday convenience*
    - Stores have the highest average workplace population within 500m
    - Typically located in major city or town centres or close by main amenities
    - Close to main local shopping and business districts
    - Have the second highest number of competing grocery retailers (reflecting their urban locations)
    - Average basket value £6.49 (standard deviation of £1.18)

- Cluster 2: *Traditional supermarket*
    - Offer the biggest range of services and largest average floorspace
    - Have the smallest average number of competitors
    - Have the lowest average residential population within 1km
    - Are typically located outside or on the outskirts of major towns
    - Sites are often near other shopping facilities and have the second highest average workplace population within 500m (this may reflect the workforce of these and surrounding stores)
    - Average basket value £18.89 (standard deviation of £5.39)

- Cluster 3: *Local convenience*
    - Located in multiple location types ranging from smaller sized town centres, rurally or close to smaller local or secondary amenities in addition to residential suburbs

- Cover a range of locations with both high and low residential and work place populations
- Average basket value £7.00 (standard deviation of £1.13)

- Cluster 4: *Student central*
  - Stores have the average highest residential population
  - Major student suburbs
  - Have the highest average number of competitors within 1km
  - Average basket value £5.46 (standard deviation of £0.30)

4.6 Cluster output discussion

For the most part the results of the cluster analysis are not unexpected, with supermarkets and convenience stores generally trading as exhibited in the previous section and being grouped accordingly. There are a few abnormalities (three supermarkets and three convenience stores), which demonstrate unexpected temporal revenue profiles affording specific insight into consumer behaviour in these stores as well as the overall insight on regular temporal patterns by cluster revenue profiles.

In the case of clusters one and three, these are predominantly made up of convenience stores (with the exception of two supermarkets in cluster one, plus one in cluster three). They all follow similar trade profiles to each other which appear more typical of convenience store trade, as identified in Figure 4.2. This classification of stores implies that the supermarkets in these clusters trade more like convenience stores in their respective cluster (i.e. stores in cluster one and stores in cluster three), following similar temporal profiles that could be the result of distinctive local temporal consumer behaviour and demand patterns, and thus are potentially operating more like a convenience store at times during the day, even though larger in size. This is supported by the lower average basket value (ABV) – the average cost in pounds for a shop at the supermarkets in this cluster, are over 70% lower than the supermarkets found in cluster 2, the 'traditional supermarket' cluster. Similarly, cluster two is almost entirely made up of supermarket stores with the exception of one convenience store. Here the cluster analysis identified that all these supermarkets followed a typical supermarket trade profile (see Figure 4.2). The exception here is one additional convenience store. This implies that this one convenience store trades more similarly to supermarkets, hour-by-hour, throughout the day. This suggests that demand for this store follows the temporal pattern of larger supermarket shopping in this location, which is likely the result of comparable temporal demand and consumer behaviour. Cluster four is made up of two convenience stores which demonstrate a much later peak in revenue, trading differently to all other stores. Both stores in cluster four are located in highly student-centred suburbs, which offer a potential explanation of this peculiar store trading pattern. This profile could thus be the result of a highly transient population, characterised by large numbers of transactions during the night and low numbers in the day, so that sales increase as students return home from university. Likewise, it could be interpreted as an indicator of distinct and inherent behaviour of student demand influenced by less structured lifestyles and different social interaction compared to typical households (Ness et al., 2002), waking up late and staying up later for example. Based on the data available in Table 4.7 it is apparent that the data partner's location type classifications, while classified according to the characteristics of each store's physical location, offer limited insight into the typical temporal revenue profiles of stores. This argument is supported by the evidence (in Table 4.7) that many location types, such as local pitch stores, are in multiple clusters. Thus, it appears that the physical characteristics of a store (as defined by the partner's traditional segmentation of location types), have less influence upon trade profiles and that temporal aspects of demand (such as demand availability) and consumer behaviour (such as how a consumer chooses to use a store to fulfil their needs as well as other potential fluctuating variables) appear more important. Instead, if we consider store formats, which are inherently different by design (or are at least intended to be) by the temporal clusters, we observe interesting patterns in temporal revenue profiles. The cluster analysis identified similarities in trade profiles across stores. These similar patterns can be used to infer the causes of revenue fluctuations by increasing our understanding of the supply of demand over time. ABV analysis of stores found that within clusters the amount spent on an average basket of groceries was similar. These values were lower in convenience stores and highest in cluster 2 – traditional supermarkets. In the instances when a supermarket was observed to follow similar temporal trade profiles throughout the day to a convenience store, the ABV value was lower supporting the assumption that consumer behaviour in these locations was different and that customers adopt a more convenience style attitude in this location, regardless of the store size. In contrast, in convenience stores found to follow a trade profile similar with supermarkets the ABV was not as high. Although temporally, sales at this location were more like a supermarket, it is likely that the limited floor space and product range inherently limit that total amount spent per shopping trip.

As demonstrated in Figure 4.11 revenue fluctuates within and between clusters over time. The shifting of demand as consumers go about their daily lives, moving expenditure from place to place, is likely to cause the periods of fluctuating revenue. While some consumer behaviour will be stochastic and therefore difficult to predict, the cluster analysis (in addition to the previous analysis in section 4.3) indicates that, in general, consumer behaviour follows repeated cycles and that for the most part consumer behaviour is habitual, exhibiting recognised and predictable patterns (Blythe, 2013, East et al., 1994, Solomon, 2013, Solomon et al., 2013). This is demonstrated by statistical similarities in the temporal revenue profiles of different stores. East et al. (1994) suggest that, in the context of grocery shopping, environmental factors affecting shopping times are largely static. This is supported by evidence of similar diurnal sales profiles observed in supermarket stores and collected over the last two decades (East et al., 1994, Ipsos-RSL, 2003). This results in regular shopping habits and regular shopping times by consumer groups accounting for these factors. The increasing availability of these data allows us to observe the major revenue fluctuations over time from real data, providing real-world insight for improving our understanding of, and making it possible to account for the impact of consumer behaviour and their movements. Using the cluster profiles presented in Figure 4.11 the main temporal fluctuations (transition points in revenue) when sustained revenue growth or decline typically occurs are observed at the following approximate times:

7am – 12:30pm (period of revenue growth – all stores)
12:30pm – 2pm (period of revenue decline – all stores)
2pm – 4/6:30pm (period of revenue growth – supermarkets/convenience)
4/6:30pm – 12am (period of revenue decline – supermarkets/convenience)

Accepting the theory that generally consumer behaviour in the grocery market follows regular shopping patterns, and that a large factor in temporal fluctuation is caused by fluctuating demand, then the transitional periods presented above (identifying specific points in the day that revenue changes) can be closely linked to specific, habitual consumer types. To explore this further it is proposed that the following demand types form examples of the separate consumer groups which when combined result in total grocery demand; resident population, workplace population, student populations, leisure population.

Of particular interest is the temporal behaviour and movements of non-residential consumers, which we note early in this thesis remain typically under-researched and are poorly accounted for in previous SIMs. Nevertheless, non-residential demand groups can be shown to contribute to grocery revenue with evidence suggesting that they do in distinctive ways, varying in terms of both behaviour and time (Berry et al., 2016, East et al., 1994, Ness et al., 2002, Newing et al., 2013b, Oh et al., 2004, Schwanen, 2004). For example, in cities where universities student numbers reach many thousands, expenditure levels are estimated at well over £100 million. Coupled with very different lifestyles and shopping behaviour from other consumer types, the impact on grocery revenue in stores supported by students is not only apparent but also produces distinctive conditions (Ness et al., 2002). This analysis suggests, for example, that this is exhibited through distinct temporal revenue profiles, such as those demonstrated in cluster four, which peak much later in the day, accounting for the different activity patterns and shopping habits of students. Similarly workplace populations have been suggested to vary significantly in shopping behaviour and shopping times from a typical residential shopper, and again represent a significant source of expenditure. For instance, research suggests that they are associated with very distinct periods of shopping, resulting in shifts in revenue occurring in the morning, leading up to lunch and early evening (Berry et al., 2016, Schwanen, 2004), with observable shifts in revenue in the partner's data seen at these times, (which Berry et al. (2016) also hint at in their research on stores in London). It is likely that in stores found to be in close proximity to high volumes of pupils at school there will also be a localised impact. From personal observation when a store is located near a school, pupils flock to it at the end of the school day to buy snacks. In the next chapter, in addition to presenting analysis of the spatiotemporal distribution of school pupils, this research seeks to examine the temporal relationship with store sales, analysing the diurnal sales profile of stores affected by school based demand, which may result in some stores demonstrating a peak in sales around 3:30-4:00pm for example.

Leisure consumers' shopping behaviour are known to vary from local/resident behaviour, demonstrating different motivations, activity planning and expenditure (Newing et al., 2013b, Timmermans et al., 2002). While it is difficult to identify customers (specifically as 'leisure' shoppers) from the data, this could suggest that this demand type is likely to be more active from mid-morning (approximately 10am) onwards**.** This mid-morning transition of activity could be the result of trip-chaining behaviour (MPTM), whereby consumers group shopping activities or other leisure activities together, which are known traits associated with consumers shopping for leisure. The purchasing of food is often perceived as a chore or necessary task and so this shopping activity is tied to more enjoyable or preferred activities or shopping, altering food shopping times to coincide within the restraints of these leisure activities. This could be linked to the periods of revenue profile activity occurring in the stores in cluster two, for example, which are mainly out-of-town, major town centre and retail park type locations which receive an increase in revenue from around 10:00-11:00am as, potentially, other stores and facilities begin to open. Explicitly linking these demand types to temporal fluctuations as well as providing accurate counts of number is problematic. However, although not represented hour by hour, research has shown that all forms of tourism (e.g. leisure) have a considerable impact on store revenue over time and that it is possible, as well as being important, to model (Newing et al., 2013a, Newing et al., 2013b, Newing et al., 2014b). It is important to estimate demand to represent these temporally distinct consumers, allocating spending accordingly in both volume and spatial distribution, as there is potential for considerable localised spatiotemporal variation to occur on a diurnal scale.

4.7 Conclusions

This chapter set out to analyse the temporal dataset in order to better understand the impact of transitional periods and how revenue in the grocery sector fluctuates over time. Data was used to provide insight into key temporal periods of revenue change, as well as highlighting distinguishable patterns, potentially affording insight on temporal demand-side components. Revenue was shown to vary over time and evidence suggests that this stems from both supply and demand-side factors (section 4.1). It was clearly demonstrated in the data that distinguishable temporal patterns of revenue exist. These patterns offer interesting and novel insights into the grocery sector, grouping stores together that were previously in separate categories when using conventional grocery sector segmentation methodologies. This made it possible to infer assumptions about consumers based on observations of supply-side temporal fluctuations. Analysis showed that while total revenue was shown to vary, in terms of volume and across different days of the week, the actual hour-by-hour temporal profiles across different days were shown to be consistent (Table 4.5 and 4.6). This temporal pattern was observed regardless of the total revenue made in stores and data showed that temporal profiles (i.e. the times when stores experienced revenue uplifts and decline) were very similar on each day of the week. This evidence led to the assumption that although consumers were shown to spend different amounts on different days (Table 4.3), hour-by-hour trade was generally periodic throughout the week, implying that consumers followed habitual shopping patterns throughout the day.

In addition to evidence of regular and habitual shopping patterns over a week, shopping times and shopping behaviour were shown to demonstrate similar characteristics and temporal trade profiles within and between different store types and locations (Figures 4.3, 4.4 4.8 and 4.9 for example). This evidence suggests that in locations which demonstrated similar temporal activity, the existence of regular and distinguishable demand types (following regular temporal patterns and exhibiting predictable behaviour) are a potentially influential factor of diurnal revenue patterns. The existence of peaks in certain locations and not in others and at certain times, led to the assertion that different consumer demand was present at these points and that demand types behave differently and that their availability appear time dependant. Based on the spatiotemporal occurrence of revenue transitions, and knowledge of the study area, potentially important demand types were identified for further analysis. For instance, early morning peaks and lunchtime peaks in the city centre were suggested to be related to workplace demand. This was corroborated with previous research, which found similar evidence suggesting the same relationship. Moreover, it was confirmed using demographic data, checking for the presence of a workplace population. The evidence suggests that certain reoccurring revenue peaks are likely to be linked to specific demand types (due to their presence in an area, as well as the behaviours they employ) and suggests that it is feasible to identify temporal demand using temporal supply-side data.

It was believed that this evidence suggested that stores with similar sales profiles therefore traded in the same way, either in terms of the way that the stores are functioning, trading more like a convenience store when it is classified as a supermarket, for example, or because consumer demand was similar during these periods in terms of time availability and behaviour. An additional observation of note, resulting from the cluster analysis, is that it appears the data partner's location types classification framework was not an appropriate categorisation of temporal activity. The evidence presented in Table 4.7 highlights how in some instances, stores of the same location type differed from one another (in terms of sales timings) and how stores with different location types, on the other hand, were similar to one another (in terms of diurnal sales patterns). Again, this evidence points to the impact of external temporal factors acting on grocery store revenue, that are independent of a store's physical characteristics, further supporting the assumption that temporal demand is important and necessary to incorporate if SIM accuracy is to be improved, as indicated by the impact upon store revenue. The theory that spatiotemporal demand fluctuations are a major factor influencing revenue will be considered further in the next chapter. The clustering framework will continue to be applied, so that it is possible to identify any evidence of similar behaviour and demand types surrounding the grouped stores, supporting this assumption as well as providing additional evidence of the source of temporal revenue fluctuations.

In the next chapter, analysis seeks to demonstrate how different demand types vary over time and space, particularly over diurnal scales, which likely contribute to the observed temporal revenue fluctuations in stores. The assumption that stores that demonstrate similar temporal profiles, and thus similar demand profiles within the cluster groups, will be addressed through further analysis of the consumer demand data. Analysis will initially focus on the partner's loyalty card data, as a source of insight into known customers and their behaviour; followed by the use of new and open source demographic data, to demonstrate spatiotemporal changes of different consumer demand. The aim is to further highlight the need for improved detail and disaggregation of temporally informed consumer demand data, when used for modelling store revenue, by demonstrating how expenditure availability fluctuates over time, and ultimately improving our understanding of spatiotemporal consumer geographies.

Chapter 5 - Identifying the spatiotemporal geography of consumers

5.1 Introduction

In the previous chapter I examined temporal fluctuations of revenue from the supply side and provided a detailed picture of evolving market conditions (based on store level revenue) that occur over time. The evidence demonstrated how varied revenue is over time. It is plausible that these variation are likely due, in part, to shifting demand and the arrival (or departure) of different demand types at certain points during the day. This chapter will now assess temporal change from a demand-side perspective, using a range of data to identify novel demand sources to use in SIMs, as well as accounting for a more complex array of behaviours. In addition, the aim is to also demonstrate the occurrence of spatiotemporal fluctuations within demand groups and the relationships between demand-side shifts with the observed temporal revenue patterns found in the supply-side data, (providing evidence of demand-side spatiotemporal fluctuations and the potential impact that this has on store revenue). The need to more accurately account for fluctuating demand resulting from spatiotemporal change represents a core goal of this thesis. This is important if SIM accuracy is to be improved through a more realistic representation of grocery demand. Following the spatiotemporal analysis of demand a SIM (utilising the temporal components of demand) will be built, incorporating novel data and behavioural mechanics, affording an increased level of accuracy and realism into store level revenue predictions.

5.2 Demand-side data: Introduction to The partner's loyalty card data

Demand-side data detailing the observed shopping behaviour of consumers has been provided in the form of the commercial partner's loyalty card dataset. The dataset contains 29 million records of average weekly sales, derived from all loyalty card customer transactions, who have purchased items in any of the partner's stores in West Yorkshire, over a 12-week period from November 2014. Table 5.1 below details the database contents. Each loyalty card customer is identifiable through an anonymised, but unique, ID number along with the output area (OA) that the card is registered in. Each item is recorded as a separate transaction alongside the product category (and value), and the commerce channel the transaction took place within. The spatial extent of the loyalty card consumer data covers all England, detailing a customers origin/'home' via OAs. However, while consumer origins are national, the store level transaction data coverage is limited to stores only within West Yorkshire.

Table 5.1 – Summary of the partner's loyalty card transaction database

Database name	Loyalty card Transactions	29,006,812 records
This table contains 12 weeks of sales data, where a loyalty card was used, ending in November 2014. The data is presented as weekly averages for the 12-week period.		
Variable name	**Description**	**Number**
CustUID	Unique ID which identifies the nectar card	918,531
_2011_OutputArea	Output Area for the customer home address.	85,540
SiteKey_As_JSBranch	ID is for the partner's store. Can be used to link to locations table.	138
CATEGORY_NAME	Category of the product purchased. Where the SUPER_CATEGORY_NAME is the hierarchal 'parent' of the CATEGORY_NAME.	238
SUPER_CATEGORY_NAME	High level group for the product, which has been purchased.	79
SupermarketItems	Number of items purchased in a supermarket.	For each relevant transaction.
SupermarketSales	Amount spent on purchases in a supermarket.	For each relevant transaction.
OnlineItems	Number of items purchased on-line.	For each relevant transaction.
OnlineSales	Amount spent on on-line purchases.	For each relevant transaction.
ConvenienceItems	Number of items purchased in a convenience store.	For each relevant transaction.
ConvenienceSales	Amount spent on convenience store purchases.	For each relevant transaction.
ShopperType	Identifies the shopper as user of convenience, on-line and supermarket stores and each combination of these.	For each relevant transaction. 5 different values.

Before any analysis was undertaken the data was processed using RStudio, and only the data relevant to the thesis was extracted, thereby reducing the volume of information. The process involved two main stages. First, OA data of consumer addresses were aggregated to the LSOA level in order to reduce spatial noise and secondly, any data detailing online sales was removed, as online sales do not fall within the agenda of this work. Processing was necessary to make sure that the dataset was compatible with the SIM used in this thesis (developed in chapter six). The consolidated data was then used to explore the geography of actual grocery store consumers below, identifying spatiotemporal demand that may explain or demonstrate relationships with the observed cluster group's temporal revenue profiles (as discussed in the previous chapter). This will be used to support the need for novel SIM extensions, justifying the need for further research into different demand types and their spatiotemporal distributions, which are suggested to represent a potential source of the temporal sales fluctuations occurring within the four cluster groups, previously identified in chapter four.

5.3 Observed consumer geography – by store clusters

The following section presents evidence of nationally observed consumer behaviour in relation to grocery shopping. The figures below demonstrate the origins of customers who shop in each cluster, as well as potential insights into grocery shopping behaviour through sales data and customer counts. The maps in the figures, presented below, were derived using observed loyalty card consumer data, with individual consumer records aggregated up to LSOA, and are used in combination with a buffer analysis (Table 5.2 below), affording additional insight in Section 5.3.2. Using these data it is possible to suggest the existence of *different* demand types (non-local/non-residential) through the spatial patterns and catchment analysis observed below, supporting evidence that residential populations alone may be a poor representation of grocery shopping behaviour and demand (a summary of the observed spatial patterns for each cluster is presented in section 5.3.2 below):

Cluster 1 – *Workday convenience*

Figure 5.1 Map showing observed sales (£) to cluster 1 stores from LSOAs

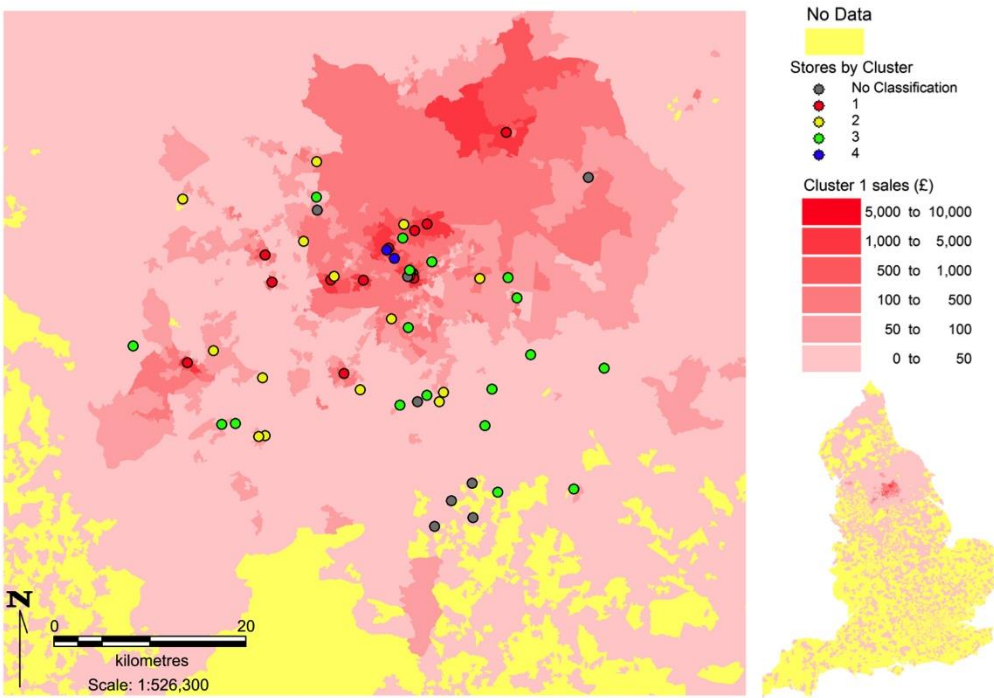
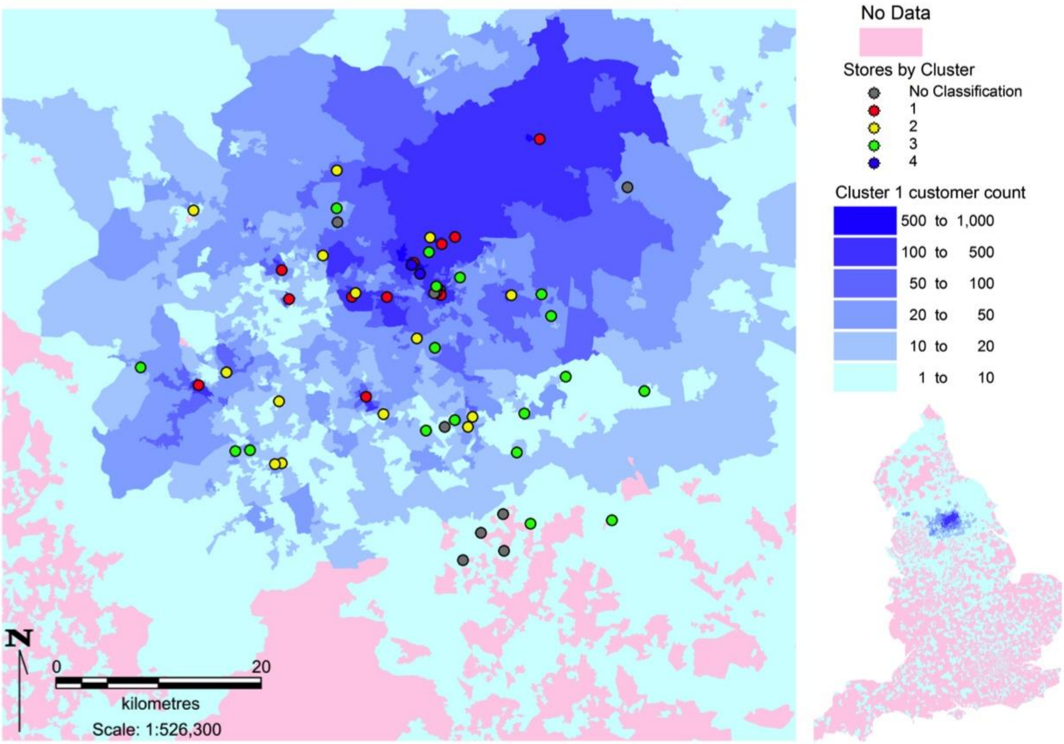


Figure 5.2 - Map showing count of unique loyalty card customers to cluster 1 stores from LSOAs



Cluster 2 – *Traditional supermarket*

Figure 5.3 - Map showing observed (£) to cluster 2 stores from LSOAs

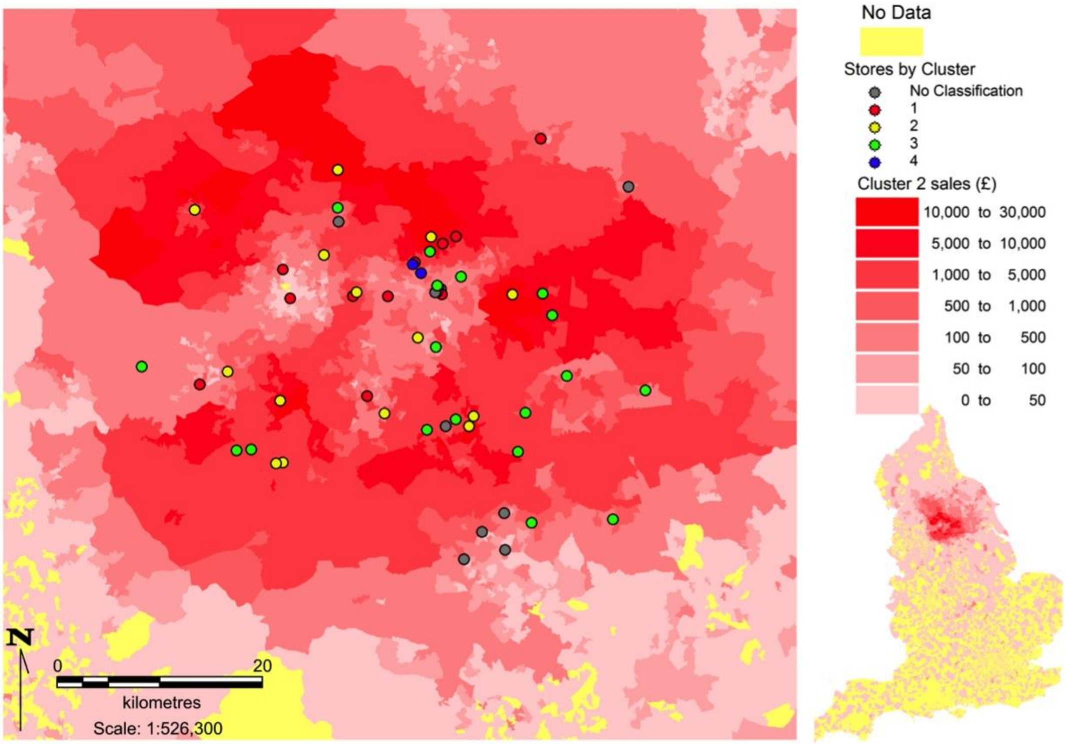
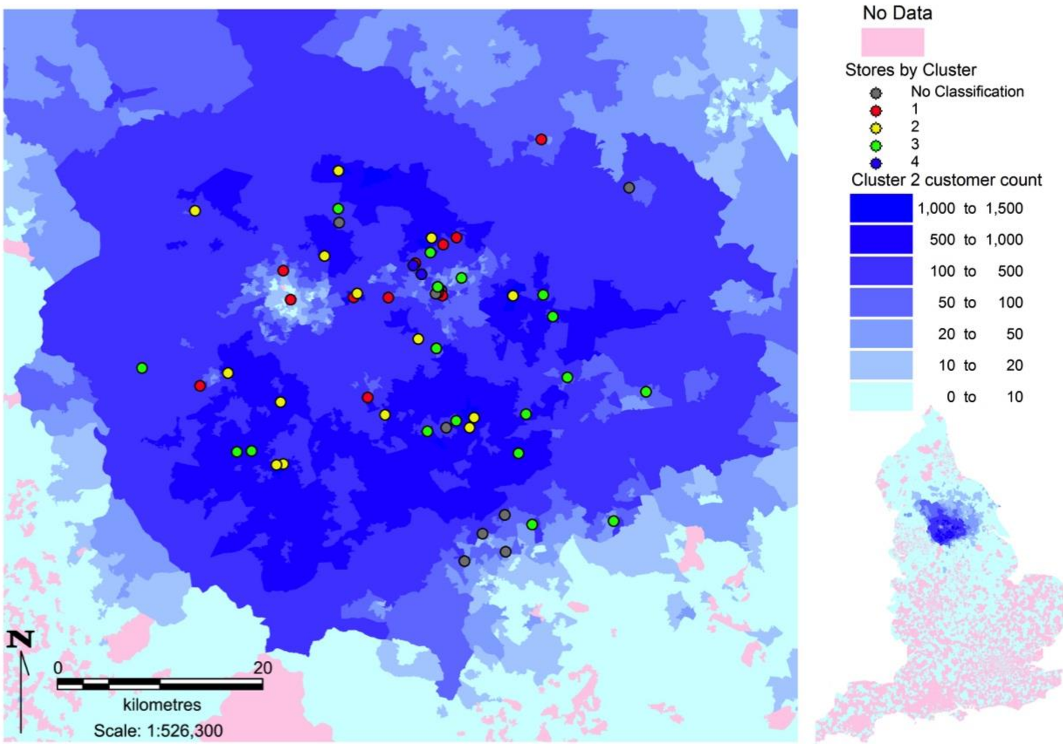


Figure 5.4 - Map showing count of unique loyalty card customers to cluster 2 stores from LSOAs



Cluster 3 - Local *convenience*

Figure 5.5 - Map showing observed sales (£) to cluster 3 stores from LSOAs

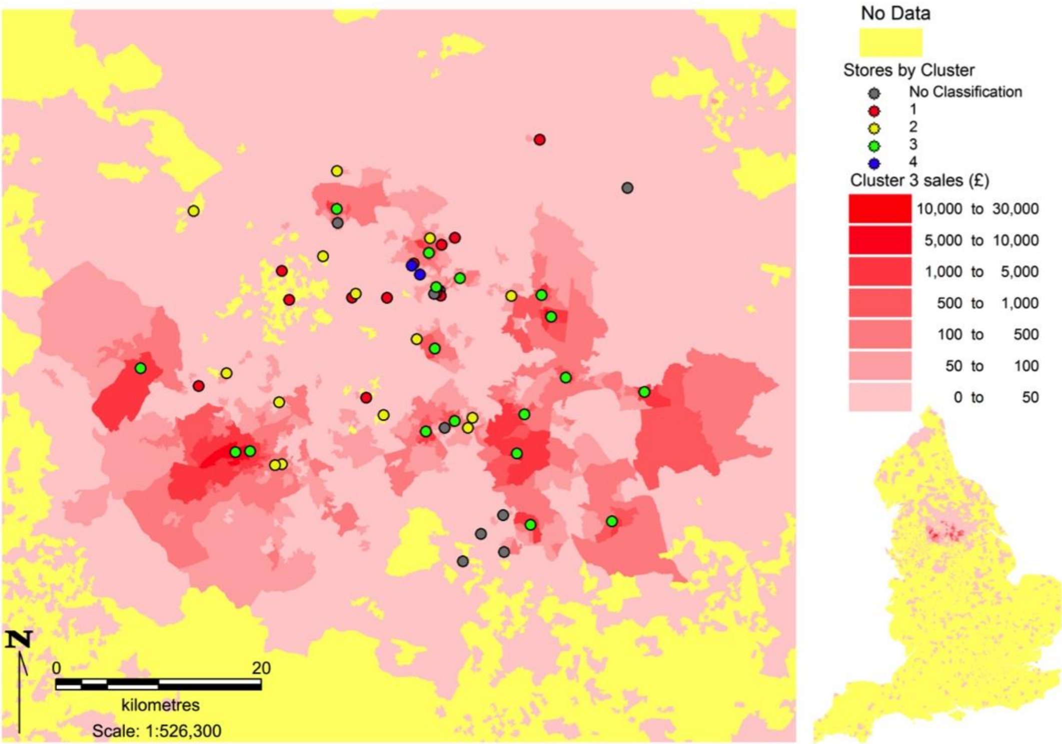
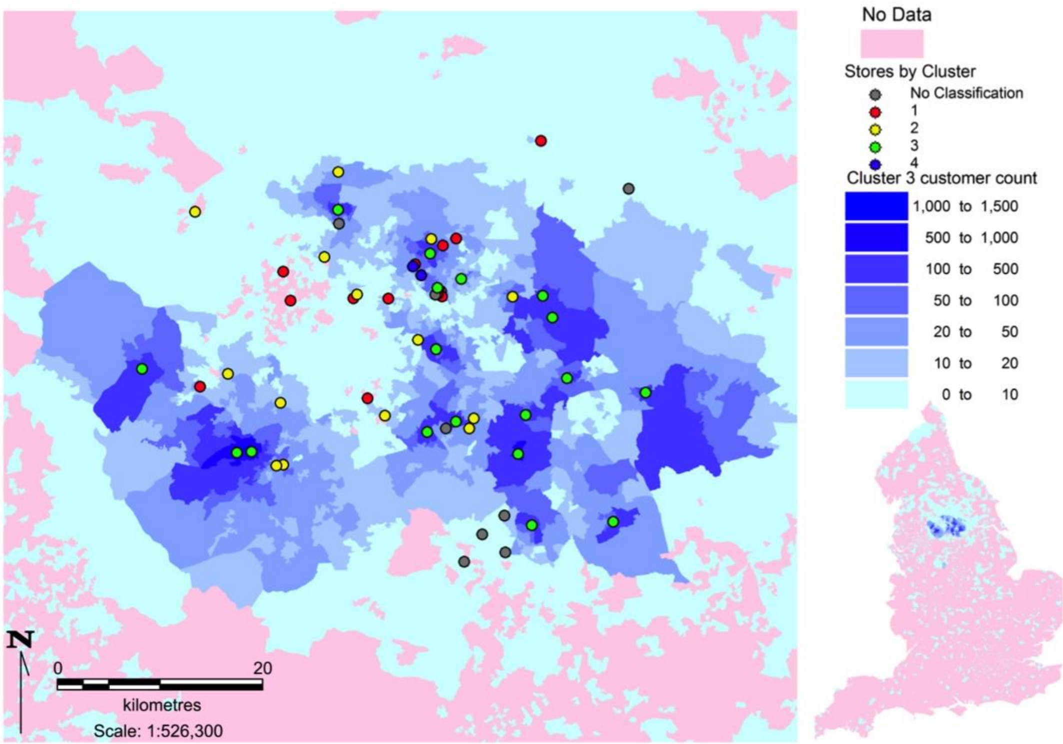


Figure 5.6 - Map showing count of unique loyalty card customers to cluster 3 stores from LSOAs



Cluster 4 – *Student central*

Figure 5.7 - Map showing observed sales (£) to cluster 4 stores from LSOAs

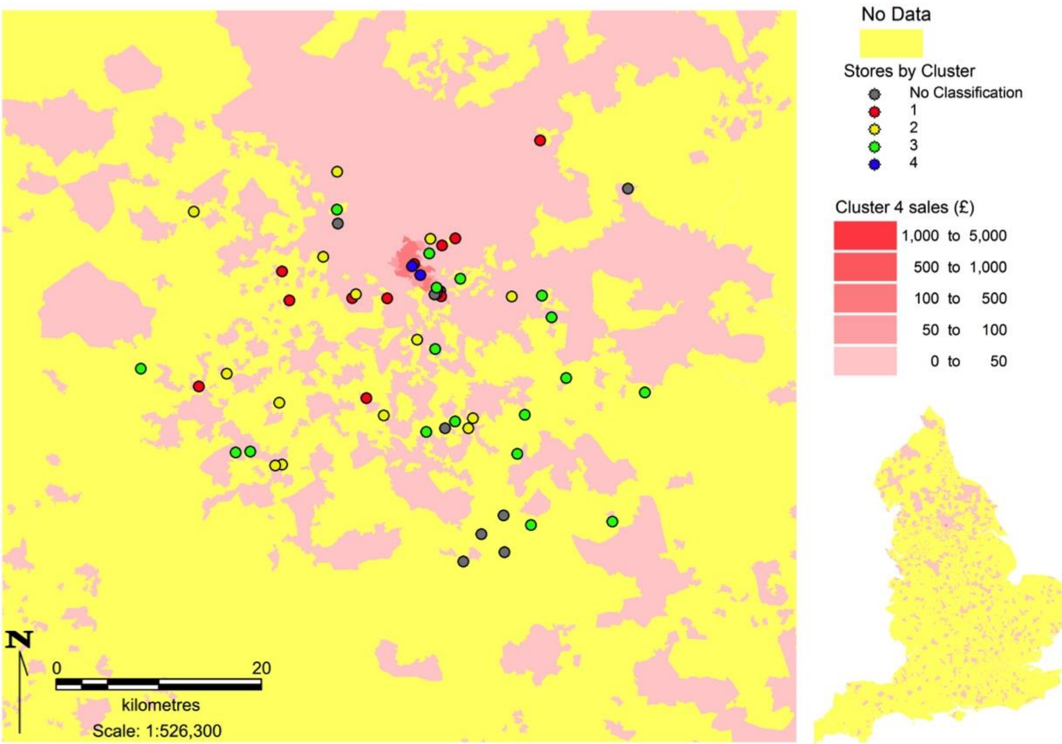
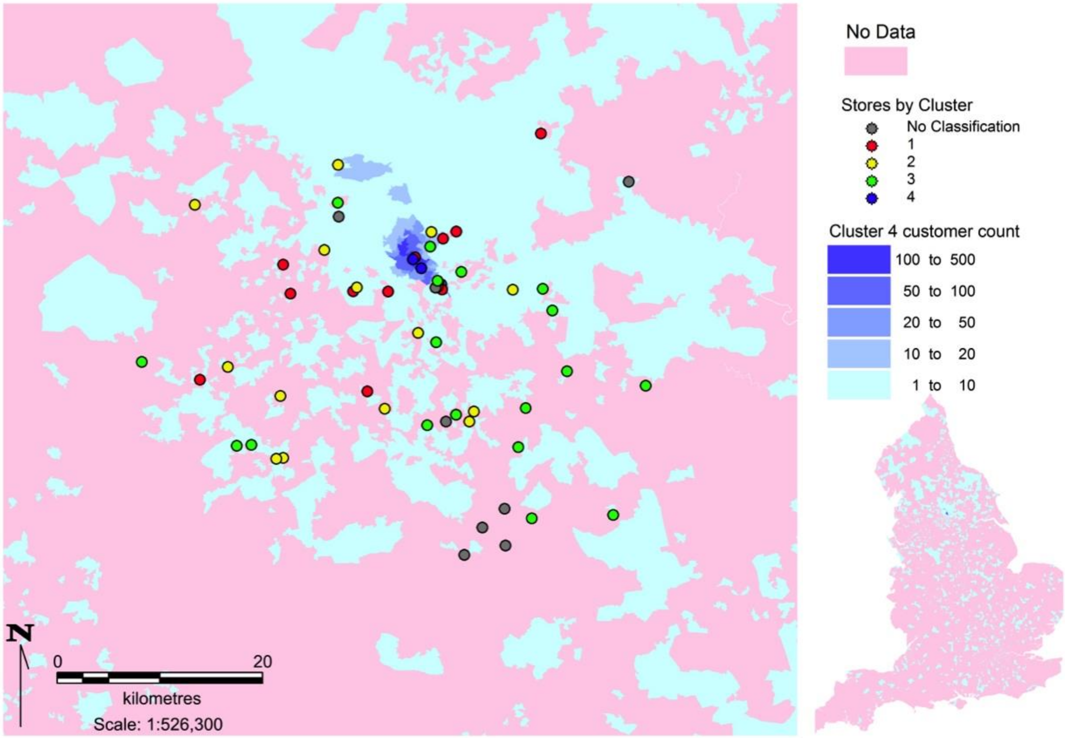


Figure 5.8 - Map showing count of unique loyalty card customers to cluster 4 stores from LSOAs



5.3.1 Demand geography analysis

Table 5.2 (below), presents the outputs of a buffer analysis on observed demand data for each cluster of stores within West Yorkshire, using national level loyalty card transaction records and LSOA level Census data. The table exhibits a range of statistics for each consecutive buffer zone around cluster group stores, providing insight into demand-side behaviour and movements, as well as identifying core catchment area sizes. A potential limitation of the data is that the loyalty card transaction records only reflect loyalty card customers, (not all grocery customers) whose behaviour may be inherently different to non-loyalty card customers as they have a tendency to be more loyal to the partner's brand. However, the lack of full customer representation by loyalty card data, in this instance, does not detract from the valuable insight of the statistical outputs, as this provides information on shopping habits, such as the uptake and shopping behaviour of specific consumer demand types, when correlated with observed demographic data in the store catchments.

** Represents the total revenue made by all stores in each cluster from loyalty card transactions.*

† Buffer distances were designed to coincide with those used in-house by the partner's location analysis team.

[1] The partner's in-house team consider a stores core catchment area to be indicated once a cumulative value of approximately 70% of revenue (£) is reached from the surrounding area.

[2] Loyalty card representation is based on the proportion of total households represented by loyalty card customers in each buffer (assuming that each household has only one registered loyalty card).

[3] WY = West Yorkshire: which is represented at LSOA geography.

[4] Represent the proportion of total residential population (derived from the census) living in each buffer zone for each cluster group.

Table 5.2 - Buffer analyses of cluster catchment areas using Loyalty card (LC) data

	Cluster group							
	One - Avg. store size 2915 (sqft) – proportion of total sales (£) made from LC 31%							
Buffer distance (km)†	Sales (£)	Customer count	% Of loyalty card total cluster revenue (£)	Cumulative %[1] of total loyalty card revenue (£)	Sales (% of total within 45km)	Customer count (% of costumers shopping in this cluster)	Proportion of total res pop[4] across clusters within WY[3] (%)	Loyalty card representation of total population within WY (%)[2]
0 - 0.5	44672	6877	15.5	15.5	17.74	9.12	1.31	57.2
0.5 - 1	71769	15144	24.9	40.3	28.51	20.09	5	33.3
1 - 2	57481	17570	19.9	60.3	22.83	23.31	14.03	13.6
2 - 5	38844	16114	13.5	**73.7**	15.43	21.37	31.08	5.6
5 -10	20047	9853	6.9	80.7	7.96	13.07	24.79	4.3
10 - 15	8582	4657	3.0	83.6	3.41	6.18	15.19	3.4
15 - 20	2473	1277	0.9	84.5	0.98	1.69	5.69	2.4
20 - 45	7896	3899	2.7	87.2	3.14	5.17	2.99	14.2
Totals	*251764*	*75391*	*288659**	-	*100*	*100*	*100*	*8.2*
	Two - Avg. store size 37858 (sqft) - proportion of total sales (£) made from LC 50%							
0 - 0.5	129580	9655	3.7	3.7	4.55	2.09	1.21	37.4
0.5 - 1	213815	24256	6.1	9.8	7.51	5.25	3.61	31.8
1 - 2	680600	78686	19.4	29.2	23.90	17.02	12.45	28.0
2 - 5	1500157	223058	42.8	**72.1**	52.69	48.26	51.44	19.5
5 -10	59855	83832	1.7	73.8	2.10	18.14	22.96	16.1
10 - 15	158221	24056	4.5	78.3	5.56	5.20	6.16	18.3
15 - 20	41249	6758	1.2	79.5	1.45	1.46	2.25	14.4
20 - 45	63743	11897	1.8	81.3	2.24	2.57	0.0	0.0
Totals	*2847220*	*462198*	*3502065**	-	*100*	*100*	*100*	*50.1*
	Three - Avg. store size 1808 (sqft) – proportion of total sales (£) made from LC 39%							
0 - 0.5	103782	11151	33.7	33.7	34.22	19.73	2.82	43.0
0.5 - 1	80047	12088	26.0	59.6	26.39	21.39	4.57	28.8
1 - 2	58463	12498	19.0	**78.6**	19.28	22.11	12.06	11.3
2 - 5	44377	14091	14.4	93.0	14.63	24.93	34.22	4.4
5 -10	11513	4400	3.7	96.7	3.80	7.78	31.59	1.5
10 - 15	2246	977	0.7	97.4	0.74	1.73	13.34	0.8
15 - 20	756	426	0.2	97.7	0.25	0.75	1.48	3.2
20 - 45	2092	888	0.7	98.4	0.69	1.57	0.0	0.0
Totals	*303276*	*56519*	*308328**	-	*100*	*100*	*100*	*6.1*
	Four - Avg. store size 1797 (sqft) – proportion of total sales (£) made from LC 15%							
0 - 0.5	4671	943	20.9	20.9	38.43	19.55	0.64	15.8
0.5 - 1	2777	1165	12.4	33.3	22.85	24.15	0.94	13.4
1 - 2	1540	759	6.9	40.2	12.67	15.73	1.98	4.2
2 - 5	1443	909	6.5	46.6	11.87	18.84	12.48	0.8
5 -10	545	394	2.4	49.1	4.48	8.17	15.93	0.3
10 - 15	296	195	1.3	50.4	2.44	4.04	22.23	0.1
15 - 20	297	144	1.3	51.7	2.44	2.99	18.31	0.1
20 - 45	584	315	2.6	54.3	4.81	6.53	27.57	0.1
Totals	*12153*	*4824*	*22368**	-	*100*	*100*	*100*	*0.5*
Total	*3414413*	*598932*	-	-	-	-	-	-

5.3.2 Summary of analysis

As shown through the analysis in section 5.3, in all cases, cluster stores demonstrate to some extent a nationally reaching customer base. While these data only represent loyalty card customers (representing a potential limitation through behavioural bias), it is likely that the more prevalent findings are indicative of the wider grocery shopping consumer population. There is evidence of customers, in all cases, who are registered from all areas of the country, suggesting that consumers are travelling greater distances than typically associated with grocery shopping. Research from the US suggested that families undertaking a grocery shop from home were reported to travel a median distance of 1.26 miles to do their shopping (Hillier et al., 2011). Similarly, space-time analysis of multi-purpose shopping (MPS) reported average distances of 1.8 and 3.9km for small and large shopping centres, respectively (Baker, 1996, 1994). The areas in close proximity to stores on the maps support this, contributing the highest volumes of sales (£). However, evidence presented in the maps (Figures 5.1 – 5.8) and Table 5.2 also demonstrates journeys far further than this, indicating that residential demand, localised around a supermarket, does not account for all grocery demand. In some cases, the observed long distance transactions could be the result of low data quality: out of date residential addresses, where customers have moved, or students for example, who may be registered at home but living away from home. However, it is unlikely this applies to all cases. The spatial patterns, as noted, appear to indicate the existence of non-local/non-residential demand, fluctuating spatially and very likely, temporally too. This demand may be the result of workers commuting into West Yorkshire, or non-local residents that live outside the core catchment but travel further, to shop at a store, as well as tourists who visit West Yorkshire for leisure. The effects of different types of spatiotemporal demand are noted by Martin et al. (2015). Their work on population modelling correlates with the observed findings in the temporal sales profiles and demand-side deductions presented here.

Following industry practice of the partner's location planning team, the core catchment area of a store is defined via the attainment of 70% of total sales, and can be derived using the areal extent at which this occurs (Dolega et al., 2016). Clusters one (*workday convenience*) and three (*local convenience*) demonstrated a core catchment size of 2km suggesting a strong link to convenience shopping behaviour and localised purchasing. Cluster two (*traditional supermarkets*), on the other hand, has a core catchment of 5km, although the sales value is proportionally quite low within 2km of the store. This suggests that customers specifically choose to purchase the groceries at this location rather than the immediate convenience of a close proximity location. The lower sales values within 2km are most likely due to the typical location of cluster two stores, such as out of town or retail park locations, which may have lower residential populations nearby. However, cluster two also demonstrates that a large proportion of its sales come from further afield, suggesting evidence of additional demand travelling to the store. Likewise, and with increased intensity the same pattern is observed with stores in cluster one, with an even greater portion of sales coming from further away, which implies why we may miss sales if only residential demand is considered. Cluster three, however, which in the previous chapter was noted to display typical convenience store characteristics, achieves 93% of total sales within 5km of the store. This suggests that in these stores, local residential demand is a major contributor to store revenue and that stores sales in this cluster are highly driven by localised demand. Cluster four (*student central*) does not reach the 70% threshold at any point in the buffering zones, suggesting that almost half of the cluster's sales come from 45+kms. The data presented in Table 5.2 demonstrates how 60% of the loyalty card sales within 45km are made within 1km of the store. It is possible that this indicates the core catchment area of stores in cluster four, with sales and the number of cluster four customers dropping off substantially after 10km. The low loyalty card representation of cluster four customers suggests that this particular demographic group has a low uptake in loyalty card programmes, but those customers who do shop at cluster four stores, are located in close proximity to the store. The visible distribution (Figure 5.7 and 5.8) of interactions with these stores, (observed to be concentrated in close proximity) supports this analysis. The area is known to have a high student population and this analysis may be indicative of student behaviour shopping traits. The data provides insight into student demand highlighting behavioural aspects such as: a low uptake in loyalty cards, but also a tendency to exhibit lower mobility, choosing to shop at the nearest available grocery store. They may also be represented by the additional loyalty card sales arriving from over 45kms away, due to having a loyalty card registered at their family home in a different address. A limitation of the loyalty card data as a device for establishing core catchment areas is that the transaction records do not directly indicate spatiotemporal fluctuations of demand. In other words, many of the sales that occur from further away could be the result of daily spatiotemporal fluctuations and therefore a customer may live far away from the store they are identified to shop in demonstrating irregular shopping habits. In reality, the transactions may be the result of the store being in close proximity to workplace or from regular trips for leisure for example. This is likely to be the case in cluster 1 figures 5.1 and 5.2 (*workday convenience*), which demonstrates flows occurring from distances not typically associated with smaller convenience store formats, and thus flows are the result of daily population movements. Consequently, it is of even greater importance for retailers to account for spatiotemporal movements of consumers as shopping behaviour (and where customers are at certain times of day in relation to their residential addresses) may vary considerably, and is observed in the data to influence store sales. Referring back to the core catchment areas derived from the loyalty card data, it is likely that in reality core catchments may be typically smaller than indicated by loyalty card sales, as sales may result from non-residential transactions originating from vary different, and potentially, much closer locations.

Nevertheless, these potentially ‘untraditional’ trip distances are again a further indicator of the importance for understanding spatiotemporal demand patterns surrounding stores.

The existence of spatially diverse, and different demand types is thus evident from the data shown in Figures 5.1-5.8 and Table 5.2, as well as the fluctuating temporal profiles discussed in the previous chapter. This again suggests, as also noted in Martin et al. (2015), the difficulty in achieving a high level of accuracy when only using residential demand for SIMs. These findings from the data represent sufficient evidence that considerable proportions of revenue are being generated by demand, which is unlikely to be traditional local residential demand. As a result, even with extensive calibration, a residentially based SIM will struggle to encompass these more diverse and fluctuating demand and behaviours.

5.4 Spatiotemporal components of grocery consumer demand

So far in this chapter, observed consumer data has been used to demonstrate the spatial extent of grocery consumers and how grocery shopping activities are not always associated with local residential populations. This has been demonstrated through GIS, (see above), and has been used in conjunction with the data analysis in chapter four to provide insight into potential demand side spatiotemporal fluctuations, resulting in different temporal revenue profiles. In this section a range of data sources, representing different consumer demand of perceived importance for the grocery sector is used to present evidence of the spatiotemporal nature of different demand types. This helps to demonstrate a source for the potential variations observed in grocery store revenue over time, which may represent a contributing factor in the fluctuating revenue profiles identified in the last chapter. This analysis seeks to generate temporally informed demand layers, which will be subsequently incorporated into a disaggregated SIM in chapter eight and represents a novel development in grocery demand and location modelling research.

5.4.1 Spatiotemporal analysis

Demand types, and a summary of their analysis, are discussed below with reference to their data sources. In addition, an analysis of their fluctuating spatiotemporal geography during a diurnal cycle is demonstrated. Modelling demand characteristics in terms of behaviour, expenditure estimates and their calibration in SIMs are presented in chapter eight.

5.4.1.1 Residential demand

Residential demand is determined via the Census of the population (ONS, 2011b). Every decade in the UK, a census of the population is conducted by the ONS in England and Wales and by the General Register Office (GRO) in Scotland. The census data used in the research has been derived from the 2011 Census of Population, the most recent full census. The measure of total population, at the time of collection, is presented via small area geography (OAs) and while this does not contain information on household income or grocery expenditure (discussed in chapter six), this data can be used to provide valuable insight into spatiotemporal distributions of residential populations. In this particular context, ultimately concerned with the impact upon the grocery industry, I distinguish population by two temporal periods: a night time population (*total residential population)* and the daytime population (*daytime residential population)*. The night time population is considered to represent the total population in their home and is a full account of the total population in any one area (Martin, 2011a, Martin, 2011b, Martin et al., 2015). It is often expressed as the night time population as it is more than likely to be one of the few times when the full population will be present in their registered residential addresses. The daytime residential population, in contrast, is defined as a count of the population who remain at home during the day (for reasons such as retirement or unemployment) and is derived by subtracting the total count of people who depart that area during the day. For example, this includes counts of workers and students, registered as employed or in secondary and higher education at the time of data collection (Martin, 2011a, Martin, 2011b). These shifting dynamics are important to accurately represent because the distinction between night time and daytime population volumes can be considerable (Martin et al., 2013). Consequently, as noted by Martin et al. (2015), a night time population is a poor representation of daytime population distributions and therefore models attempting to predict store revenue will be limited in accuracy due to the observed conditions potentially being considerably different at different times of day. The number and type of people in different locations will have a profound impact on a store's revenue, resulting in for example, the fluctuating temporal revenue profiles presented in the previous chapter.

Below, Figures 5.9 and 5.10 represent the residential population (typically used for modelling) and the level of redistribution which would result in a distinctly different residential daytime population. The data presented in Figure 5.10 illustrates the level of temporal change in residential population counts on a daily basis, which is then redistributed throughout the region. The values include the count of the population leaving an area during the day or those who are classified as different demand groups (See section 5.4.1.4), which in this case is due to distinct differences in perceived shopping behaviour (Ness et al., 2002). A particularly affected area by temporal demand, through a significant drop in residential population numbers occurs in Leeds. Several LSOAs have a considerable reduction in typical residential populations, due to such high numbers of university students living in those LSOAs.

Figure 5.9 - Total residential populations by LSOAs in West Yorkshire

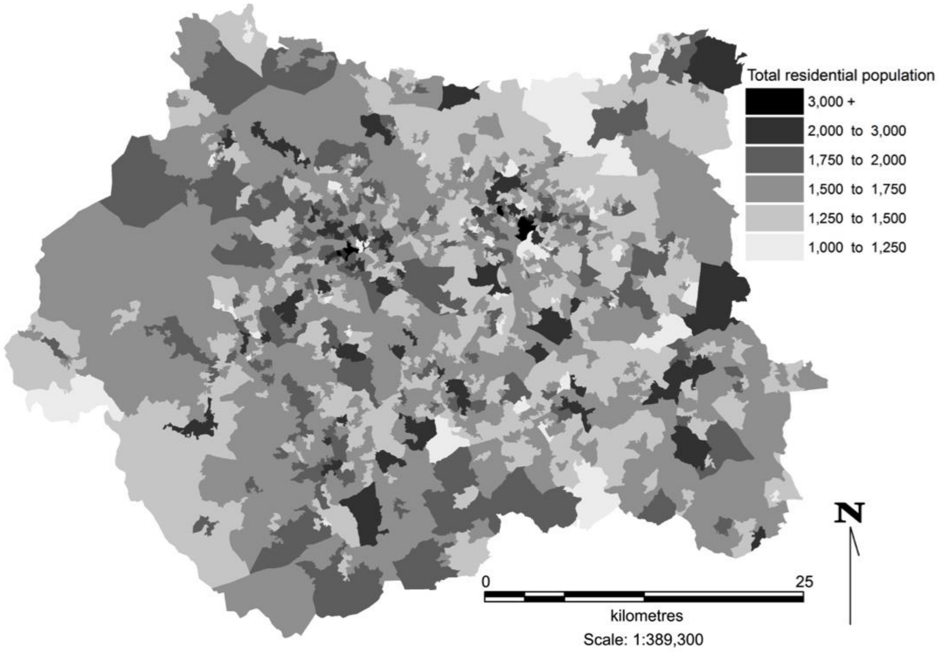
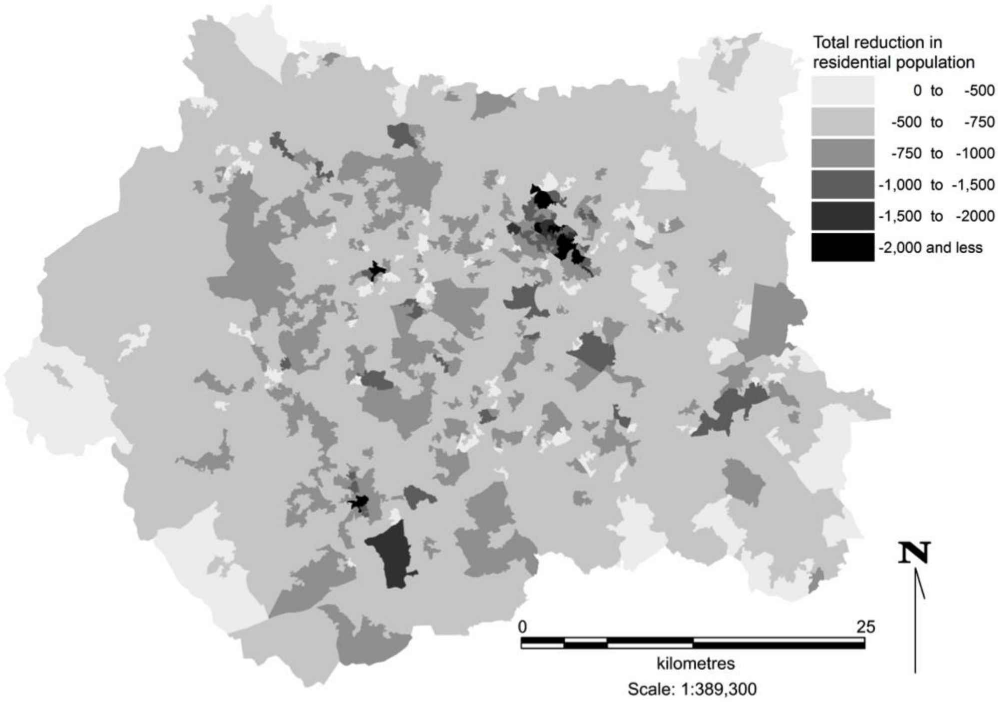


Figure 5.10 - Total counts of population reductions resulting from spatiotemporal fluctuations during the day and use of a more diverse set of demand types by LSOA



5.4.1.2 Workplace demand

Workplace demand is derived from census workplace population statistics which have been designed to supplement residential data, based on the workplace population (Mitchell, 2014). The new workplace zone geography (WPZs) was created to provide detailed counts of the workplace population at their usual place of work across the UK. While WPZs were not designed to coincide with LSOAs, they are nested within MSOAs and therefore a comparison of spatiotemporal patterns in terms of population movements within West Yorkshire can be visualised (Figures 5.11 and 5.12). However, the two ONS geographies (LSOAs and WPZs) are distinctly different, particularly in city centres with low residential and high workplace populations with many LSOAs merged together. A fully detailed explanation of the workplace data and WPZs can be found in Chapter seven. The purpose of this section is to demonstrate the extensive spatiotemporal changes that take place in many areas of West Yorkshire, when transitioning from night to day due to the working population. The total employed residential population in West Yorkshire, derived from the census, is approximately 780,000 (shown in Figure 5.11), whereas the total workplace population, derived from workplace statistics, totals over 1 million individuals (shown in Figure 5.12). One particular LSOA in Leeds city centre goes from 424 employed residents to 32500 workers, and is represented by 59 individual WPZs. The analysis reveals a considerable transition in workplace demand, not only in terms of total numbers but also in terms of the spatial distribution around the area. The observed trend for workplace demand tends to be a flow of demand from more rural/suburban residential areas towards larger more developed urban centres over this time period.

Figure 5.11 - Total counts of employed individuals at their usual place of residence by LSOAs

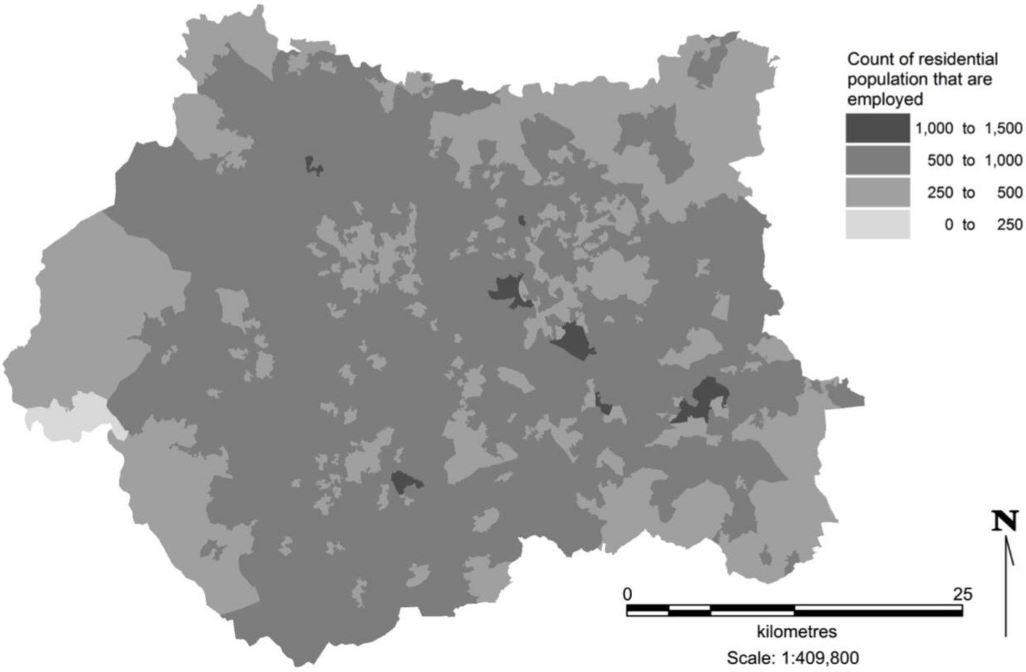
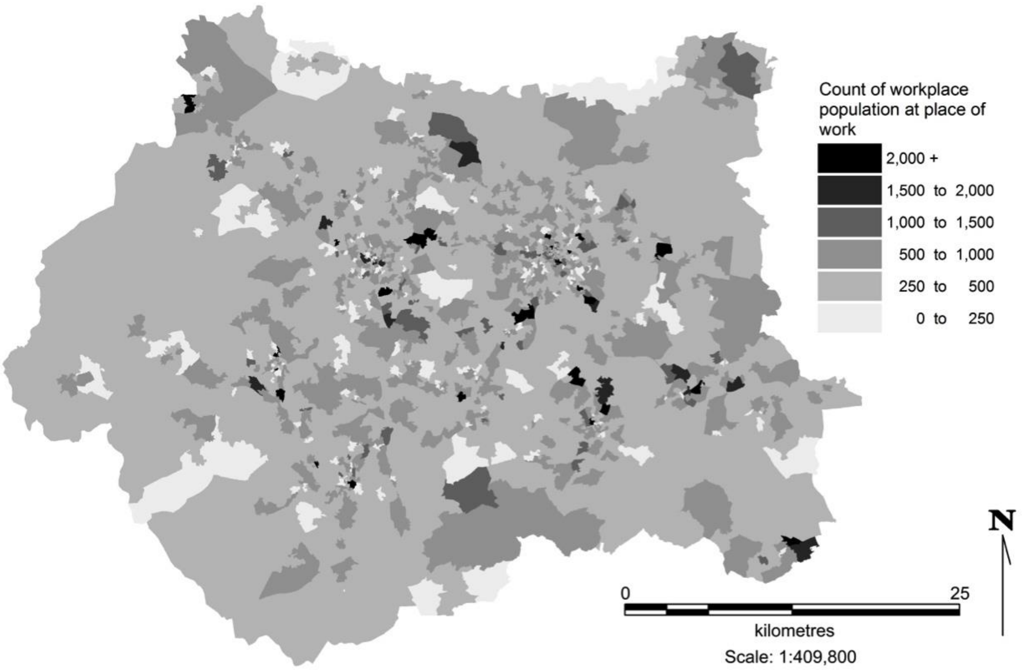


Figure 5.12 - Total counts of the working population in their usual place of work by WPZs



5.4.1.3 School based demand

School based demand is derived from two data sources: the count of the total residential population in the secondary school education age bracket and their home location as derived from the 2011 Census. This was calculated by taking the total count of children aged between

11-15, who are in compulsory education and the total count of 16-17 year olds who reported themselves, in the census, as full time students. The daytime count of school-based demand is represented by the total number of secondary school students enrolled at each school in each LSOA in West Yorkshire. The data was obtained from the Department for Education's Schools, pupils and their characteristics database (Department of Education, 2015) and was spatially referenced. The decision to exclude primary school children from a school based demand was based on the premise that primary school children are unlikely to patronise grocery stores without other family members and therefore, spending is already being accounted for via household expenditure estimates. Research notes that adolescents, on the other hand, are more likely to consume fewer meals with their family than younger children, consuming only 65% of their total energy intake, with the rest being made up elsewhere and increasingly by unhealthy food (Larson et al., 2006, Powell et al., 2007). The maps, presented below, show the total count of secondary school age children by LSOAs in West Yorkshire (Figure 5.13) and the total count of secondary school age pupils enrolled in schools in West Yorkshire (Figure 5.14). The observed spatial patterns are directly linked to the location of schools. It is possible to make realistic assumptions regarding the spatial and temporal side of school based demand movements based on the typical state school timetable. One observation, on the spatial transition of school-based demand during the day of interest, is that demand becomes far more concentrated, grouping adolescent consumers together. This is likely to have a notable impact on any retailers in the area due to high concentrations of consumers in a single location, who are highly time restricted and who are noted to more than likely engage in similar shopping activities, and to do so together, (Beatty et al., 2015, Gentina and Chandon, 2013) and thus, will have some uplift in sales as a result at certain times of day.

Figure 5.13 - Total counts of secondary school age of children in education in each West Yorkshire LSOA

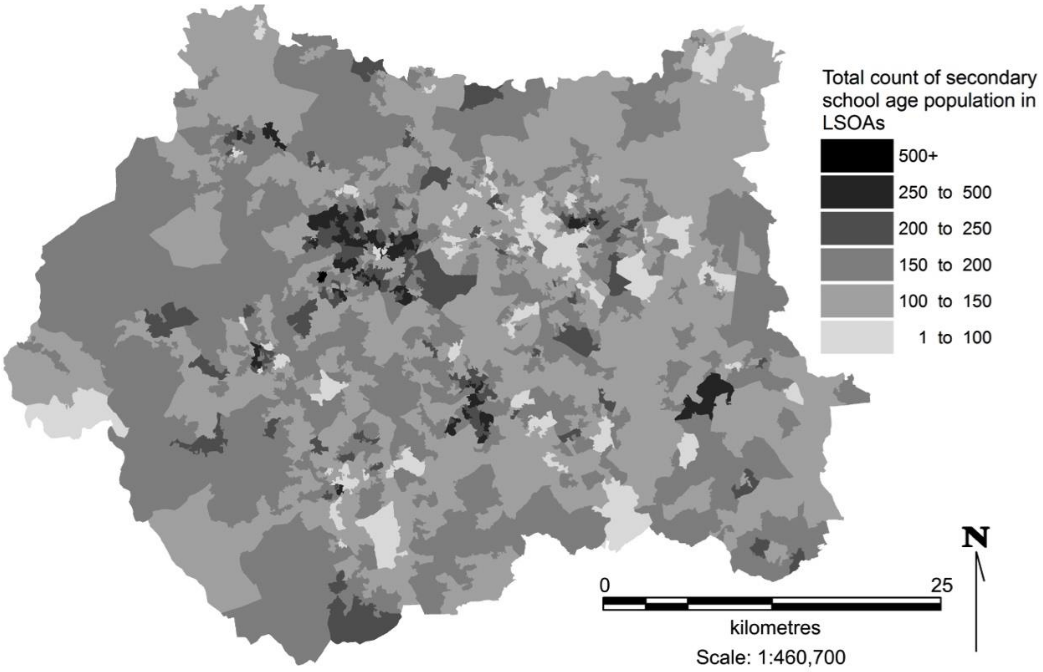
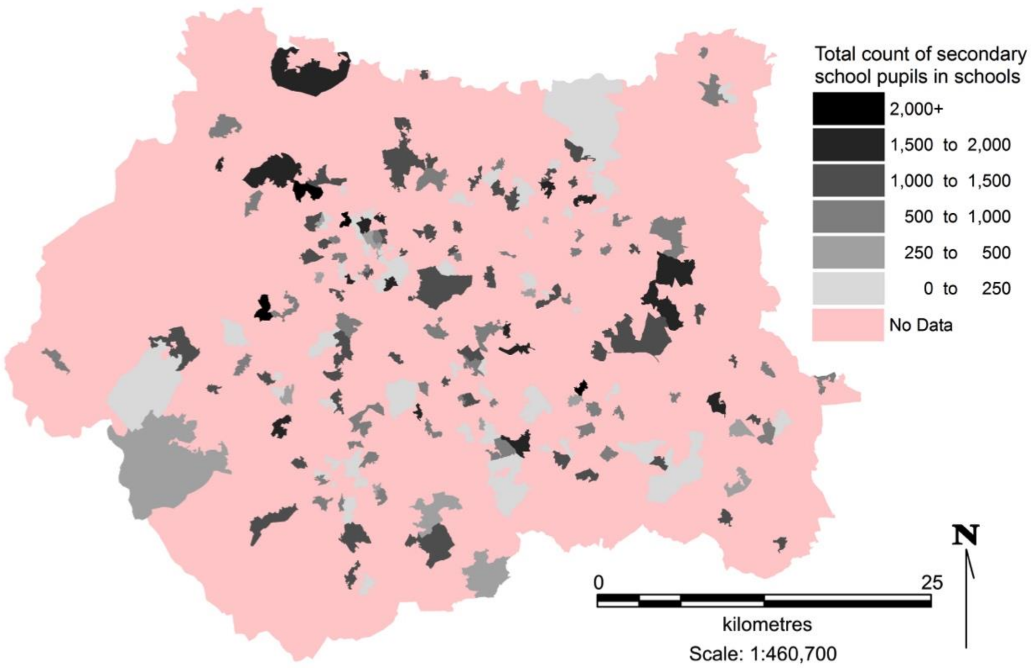


Figure 5.14 - Total counts of secondary school age pupils in schools in West Yorkshire by LSOAs



5.4.1.4 University student demand

University student demand has been estimated from two sources: first the usual place of residence of students ascertained using the 2011 census. The data used indicate a total count of the usual residential population of 18+ year olds who declared themselves as students. The data accounts for students who live at home even during term time, as well as those who live at a term time address, classifying the term time address as the usual place of residence when no alternative home address was provided in the census. This is represented at LSOA level (ONS, 2011b, ONS, 2014a, ONS, 2014b). Second, data providing counts of student enrolment at universities throughout the study region was derived from the Higher Education Statistics Agency (HESA). This provides total counts of higher education students, by higher education provider using the university's postal address as a spatial reference for the daytime location of students. A limitation of this data is that by representing counts of university level students at the postal address it does not account for large or separate campus locations, nor does it account for students' timetables and irregular contact hours. The result is that a highly concentrated population may be applied to a single LSOA and in far greater numbers than those that occur in reality. This is important to address, particularly when modelling retail behaviour as it impacts expenditure distributions and demand availability. This may have substantial impact upon grocery market retailing and in modelling store revenue. However, in order to limit this, the following rule has been applied for daytime distributions of university demand: a 10% (at home) and 90% (on campus) split. This framework has been adapted from the findings of Edwards and Bell (2013), who suggest that, an approximate figure of 90% of total university population is found on campus at midday, following the results of a survey of campus populations. Figure 5.15 demonstrates a count of the usual university age student residential population derived from the 2011 census and Figure 5.16 demonstrates the total count of all enrolled students at higher education providers in West Yorkshire and those at home during the day based on the 90/10% split. Figure 5.16 demonstrates the potential student populations at any of the university campuses. It is highly unlikely that this maximum capacity is reached at any single point due to irregular time constraints on individual students and this is represented via the portion of demand at home during the day (Charles-Edwards and Bell, 2013, Tomlinson et al., 1973). While evidence suggests university demand exhibits a series of distinctive behaviours from other demand types making them identifiable, their temporal dynamics are more complex due to the semi-unstructured nature of daily activities and lifestyles and it is noted that there is much diversity even within a similarly behaving population (Charles-Edwards and Bell, 2013, Ness et al., 2002, Tomlinson et al., 1973). The decision to model university student demand separately is based on evidence provided by Ness et al. (2002), which suggests that their behaviour is inherently different to typical grocery shopping populations. It is likely that in addressing university demand, and accounting for their unique and complex grocery shopping behaviour and their expenditure estimates, (in addition to applying a temporal framework to account for the temporal fluctuations exhibited by university demand) SIM accuracy will improve. This is discussed further in chapter eight where an estimate of a temporally fluctuating university demand is implemented into a SIM.

Figure 5.15 - Total counts of usual residential university age student population by LSOAs and main West Yorkshire university campus locations

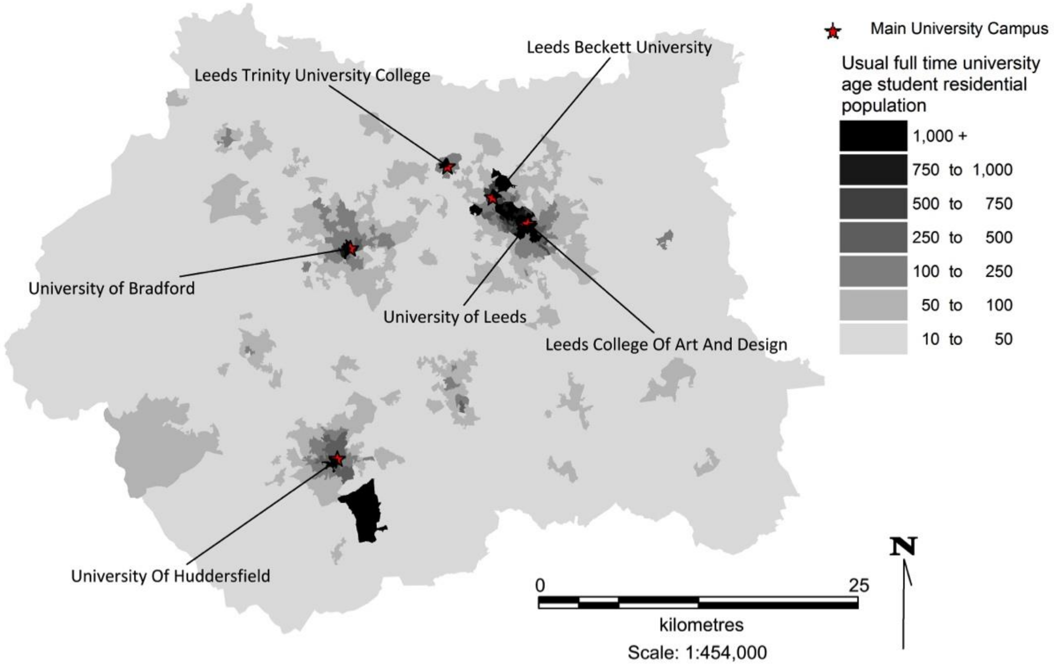
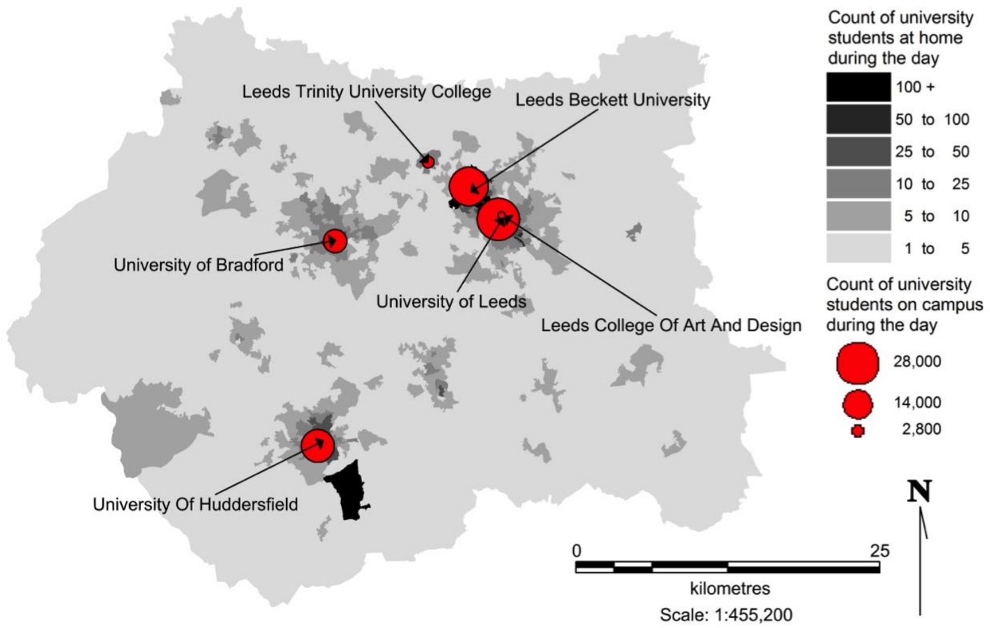


Figure 5.16 - Estimated counts of university demand at home and on campus during the day in West Yorkshire, using a 90/10% split



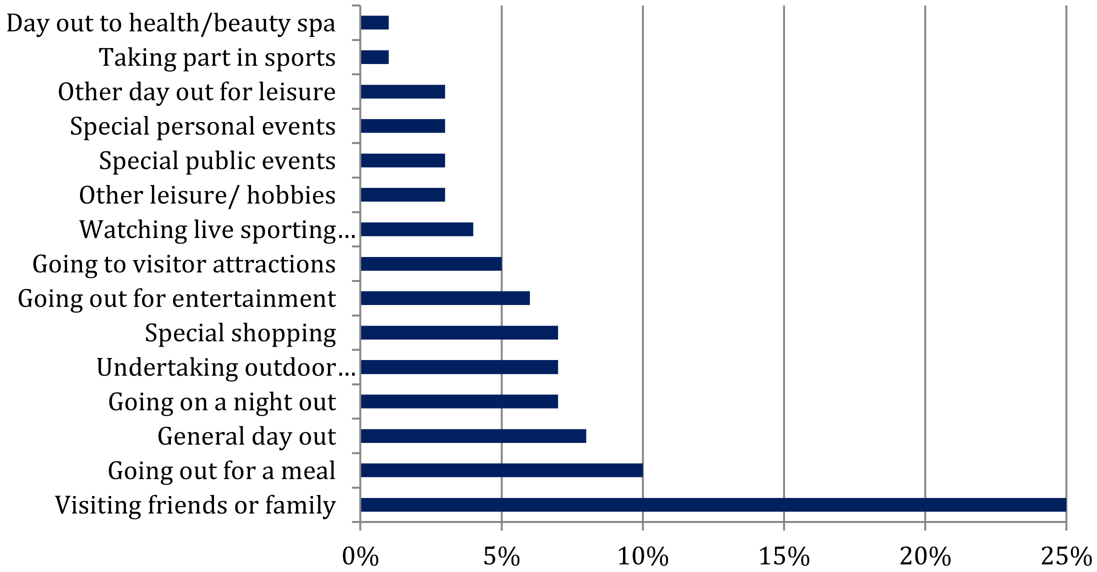
5.4.1.5 Leisure demand

Tourism demand is a growing component of our present day society, contributing vast quantities of expenditure into the economy (Au and Law, 2002, Song and Li, 2008). The data represented below, as an indicator of leisure demand, has been derived from The Great Britain Day Visits survey 2015. Visitor volumes are derived from Tourism Day Visits, which are defined as lasting longer than 3 hours, not undertaken regularly, at a destination different to the origin and involving at least one of the activities shown below in Figure 5.17 (Orrell et al., 2015). While eating out represents the second biggest activity for days out in Great Britain, this relates to food and drink purchased in cafes, restaurants and other similar establishments. This does not account for food and drink bought at retailers. It has been claimed that grocery purchases made by tourists represent such a small proportion of grocery sales that groceries therefore can be seen not as a tourism commodity (Wilton, 2006). However, more recent studies into tourism demand dispute this claim (Newing et al., 2013a), claiming that tourism demand has an impact on grocery sales. In particular it is noted that in holiday destinations during the summer months, revenue generated by tourism will supplement residential sales at supermarket stores for instance (Dudding and Ryan, 2000). This certainty is supported by the operational actions of several grocery suppliers in the UK, noting the potential expenditure to be gained from leisure demand. Spar, for instance have opened several stores at holiday park sites across the UK, supplying to visitor demand (Briggs, 2010). Likewise, Waitrose ran marketing campaigns promoting online grocery shopping deliveries to tourist accommodation, which Tesco are still actively offering (Newing et al., 2013a, Tesco, 2016). While in many cases tourists may bring supplies with them, or as discussed above eat out, some spending in food and drink retailers will take place at the destination, with grocery retailers providing these vital services to tourists (Newing et al., 2013a, Timothy, 2005).

Table 5.3 - Average number of visitors and expenditure to West Yorkshire LADs (*in Millions*) - adapted from (Orrell et al., 2015) p90.

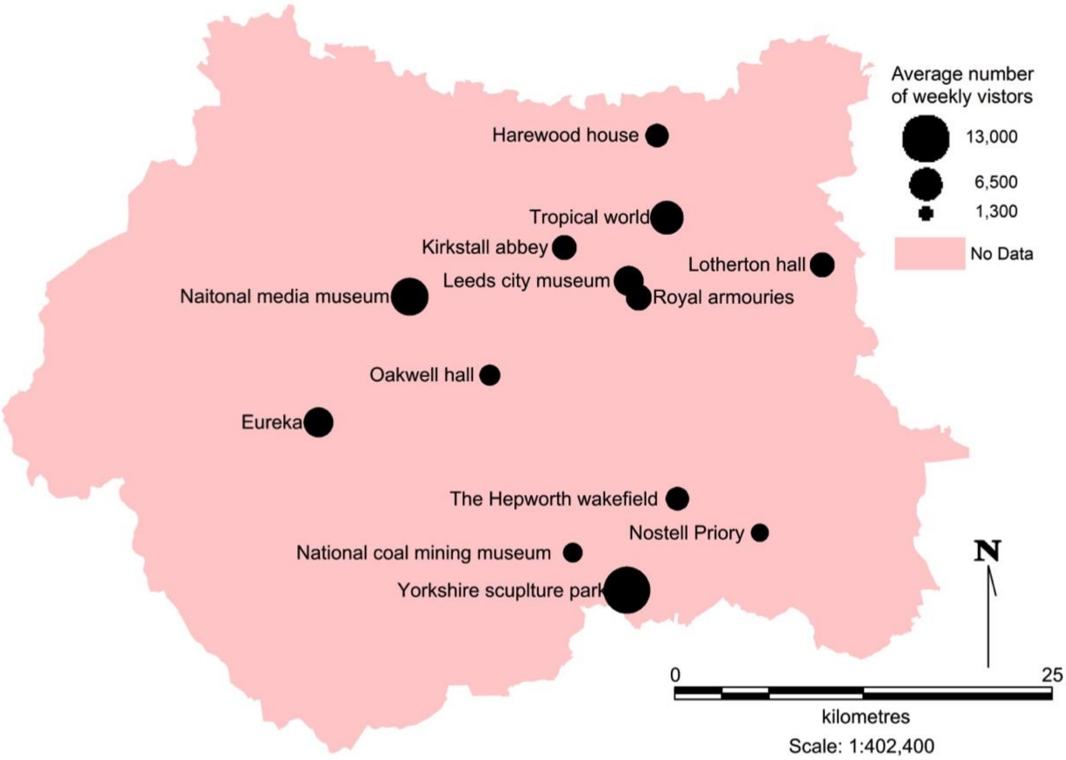
LAD	Avg. annual visitors	Avg. annual expenditure (£)	Avg. weekly visitors	Avg. weekly expenditure (£)
Leeds	21.7	737	0.42	14.2
Kirklees	9.7	246	0.19	4.7
Bradford	7.6	178	0.15	3.4
Calderdale	4.6	162	0.09	3.1
Wakefield	6.07	138	0.12	2.7

Figure 5.17 - Volume of visits by main activity type for all GB residents (% of total) – adapted from (Orrell et al., 2015) p15.



Although increases in store revenue have been suggested to be linked to tourism through increasing numbers of visitors to an area, it is difficult to correlate exact sales figures with tourism demand as data is difficult to obtain, affording insufficient clarity (Newing et al., 2013a). In Yorkshire and Humberside the average distance travelled by visitors on day visits was 54 miles (Orrell et al., 2015). Considering the buffer analysis in Table 5.2, although only a small fraction of the sales, a small proportion of revenue was observed to originate from a distance of over 45 km (28 miles). This evidence implies, similarly to the suggestions presented in Newing et al. (2013), that customers are travelling into the area and this has some level of impact upon grocery store sales that may be the result of tourism demand. It is unknown and difficult to categorically state whether these loyalty card sales originating from these areas are linked to tourism and not for instance the result of customers commuting to work. It may be possible to indicate this via the frequency of visits made by a particular customer, revealing the regularity of trips, if access to this data was available. This level of detail was not available from the loyalty card dataset, which recorded individual products as a single transaction and not as individual baskets. Nonetheless it is important to consider, and to at least demonstrate, the fluctuation of leisure demand and the level of spatiotemporal change that visitors demonstrate. The average numbers of weekly visitors to West Yorkshire represent some of the highest values of visitors found in all English LADs (total West Yorkshire values shown above in Table 5.3). However, lower level geography data is sparse and although 'going to attractions' represents only 5% of GB visits (Orrell et al., 2015), some attraction specific data was obtained to demonstrate the volumes of additional demand arriving in areas effected by visitors and is demonstrated below in Figure 5.18. This highlights the average number of weekly visitors to attractions. The data represents a count of visits to the top twenty, paid and free, visitor attractions found in West Yorkshire (Visit England, 2015a, 2015b), indicating additional weekly demand arriving in these locations. Until relatively recently, despite links to sales fluctuations (Newing et al., 2013a), tourism demand has received surprisingly little attention in SIMs for the grocery sector. Additional leisure demand, which is often spatially constrained by activity types and by destination and is typically temporally fluctuating (diurnally and seasonally), may well represent a missing proportion of revenue in SIMs and is useful to consider in future revenue predictions.

Figure 5.18 - Count of average weekly visits to the most popular free and paid visitor attractions in Yorkshire and Humberside for West Yorkshire



5.4.2 Cluster level fluctuation

Tables 5.4 and 5.5 (shown below) represent counts of, and the proportion of, various demand groups within incremental buffer zones (km) around each of the store cluster type groups for night and daytime periods respectively. In both instances, different types of demand are demonstrated, alongside traditional demand data in the form of total local residential population counts. This is to demonstrate the core themes of the research: the first being that a count of total residential population alone is a too simplistic representation of customers and that in reality there is a much more diverse customer basis within catchments which will likewise have very different behaviours. Secondly, that a count of total residential population, which typically represents a sedentary population at home over night, does not accurately represent changes that occur due to temporal fluctuations of catchment populations particularly during the day. This supports the aims of this thesis: to develop a series of temporally informed, more realistic series of demand layers that incorporate different subsets of consumer population. Before addressing the insight that the buffer analysis (below) reveals, it is necessary to first explain the two tables below so that the evidence provides a clear and concise argument.

Table 5.4 demonstrates the total count of residential populations at night by cluster types within certain buffer distances. The traditional residential count accounts for all of the registered population living in each buffer zone and represents their registered address and the likely location that they would reside in overnight. Total residential demand is then a subset of four demand types, with different temporal patterns, behaviour and lifestyles and exhibits different interactions with grocery retail. While the traditional residential count equals the same volume of demand of each of the disaggregated demand types combined, it does not account for the different aspects of consumer behaviour. Each of the demand types (as defined above) is represented by a count of their respective population in each buffer zone. Each demand type is accompanied with a percentage value of their presence within each buffer zone as well as a total presence within the overall core catchment of each cluster of stores.

Table 5.5 demonstrates the total count of daytime populations within each buffer zone, again shown by the traditional residential population count and by disaggregated demand types. This time population count is based on the location that that demand resides in during the day, which (with the addition of leisure demand, caused by the arrival of visitors at the main attractions throughout West Yorkshire) provides an estimate of the total daytime population found in each buffer zone. The data in Table 5.5 offers insight into why revenue predictions may differ from observed sales as simulated populations using only residential demand are unrealistic and fail to incorporate spatiotemporal fluctuations as customer demand [expenditure] increases and decreases in an area. Again, the demand types are as defined above. However to reiterate, residential daytime demand refers to individuals who are unemployed/retired, are not students, are off more that day and are likely to be at their home during the day. University demand is a count of full time students over 18: this is split during the day between their homes and campuses based on a 10/90% split. This division, as noted above, is based on the research of Edwards and Bell (2013), whose research suggests that 90% of a university's population are on campus at midday. School based demand is a total count of secondary school pupil numbers by each school and is also spatially referenced. Workplace demand is a count of the workforce at their usual place of work and leisure demand is a count of day visitors to the top 20 (paid or free) attractions in West Yorkshire found in the study area, based on weekly average visitors.

At a cluster level the observed fluctuations in demand (as catchment populations transition from night time in Table 5.4 to the day time in Table 5.5) can be used to strengthen previous assumptions about spatiotemporally fluctuating grocery demand. Cluster one stores demonstrate a marked increase in the total demand within the core catchment of these stores during the day. This is particularly observed in the changes to the volume of residents and workplace demand at their place of work. This cluster type, previously described as *workday convenience*, is predominately served by a temporal workplace demand. The fact that a traditional population count only represents 85% of the total daytime population highlights the flow of people into the catchment during the day. The assumption that the revenue profile is influenced by temporal demand is also supported by the fact that 67% of demand within 500m of the stores is workplace, which has previously been suggested as the typical distance travelled by workers when purchasing food and drink (Berry et al., 2016). The timings of increased and reduced observed revenue (Figure 4.11) over time are indicative of the previously suggested behavioural aspects of workplace demand. Individuals might shop on the way to and from work and during their lunch breaks, shop within close proximity to their workplace and for a short period of time (Berry et al., 2016, Schwanen, 2004), further supporting the need to account for temporal fluctuation.

Cluster two, previously described as *traditional supermarket* based on their characteristics and the observed temporal revenue profiles, are equally supported by residential and workplace demand during the day. Similar to cluster one, the total volume of demand increases in the core catchment for cluster two during the day. As time progresses through the proportion of local residents decreases being replaced with an increased proportion of workplace demand in the area. This temporal dynamic, regarding daytime populations influenced by residential and workplace demand, may reflect what could be described as a traditional supermarket revenue profile. This is observed in cluster two and appears to be the general scenario for supermarket stores and is likely to be easier to predict. The influence of local residential demand is supported by the findings of East et al. (1993), who noted that residential behaviour at supermarkets followed a more traditional framework of shopping (as described in Chapter two), noting that a similar temporal pattern regarding shopping trips in supermarkets occurred throughout the day. Likewise, the observed increase in workplace numbers and proportional share of demand, observed in Table 5.5, may explain why a series of [small] peaks are observed at similar times of day to those found in the workday convenience cluster at midday and in the early evening (see Figure 4.11).

Cluster three has been termed *local convenience* due to the fact that stores in this cluster exhibited an observed revenue profile (in addition to store location and store level characteristics) of stores that are potentially influenced by both residential and workplace demand at periods throughout the day, an assumption that is supported by the evidence of temporal demand fluctuation shown in cluster three's core catchment in Tables 5.4 and 5.5. The data suggests that although there is still a considerable volume of residents at home during the day, a large proportion of demand leaves the area with the total residential population almost halving. However, while the number has decreased, the daytime residential population surrounding these store types is the highest across all clusters, clearly indicating that residential demand in areas surrounding these stores is still very prevalent and may offer an explanation of the generally high local trade. Similarly, workplace demand is observed to almost double in volume during the day and it gains a notably increased share of the core catchment demand.

Cluster four, which was coined *student central,* has a considerable transition, (particularly in very close proximity to the stores) between the night time and daytime population. The stores in this cluster are highly dependent on university demand on a night time, making up over half of the total demand within 1 km of the store. Based on the concentration of sales, and customer counts from loyalty card data, it is likely that the core catchment for stores in this cluster is comparatively smaller than the other clusters (Figures 4.7 and 4.8). Subsequently, it is likely that the behaviour of university consumers, which has been noted to be distinctive even to their non-student peers (Ness et al., 2002), is highly sensitive to distance and travel costs. The result is that university demand is highly immobile with students choosing to shop at the stores in closest proximity. Therefore a shift in the population, such as the major proportions of the population from their homes to the campus as demonstrated in Tables 5.4 and 5.5, has a great impact on the total demand available for stores in this cluster during the day. Their absence during the daytime and their return en-masse during the evening, as well as their limited mobility, may explain why the temporal profile of these stores is different from the other clusters and why revenue instead peaks during the night (Figure 4.11). Based on evidence from the observed loyalty card data and the temporal revenue profile, one could suggest that university students are less likely to travel further than 500 metres to purchase basic groceries. This assumption is based on evidence in Table 5.5 demonstrating that during the day 60% of the core catchment demand is represented by students on campus. However, the high presence of this particular demand type does not appear to be reflected at the same time of day (in the store revenue sales profile) for that part of the day, indicating that they have a low mobility. Therefore revenue in stores highly influenced by university demand will largely reflect the spatiotemporal fluctuations of university students and their grocery shopping habits.

Overall leisure and school-based demand appear to have limited presence in any of the cluster buffer zones, though this is only representative of the data partner's stores. However, as posited in the previous chapter, I suggest that at a store level the impact of these demand types felt at certain times during the day will likely be experienced in individual stores and while proportionally they may contribute to only a small portion of overall demand, the local impacts will still be present. For instance, analysis of two of the partner's stores found in immediate proximity to schools demonstrated a sharp uplift in sales at 1530, typically representing the end of the school day. It is not possible to directly relate this uplift in sales to schools. However, two stores found furthest away from schools appeared to demonstrate no evidence of a similar peak in sales at this time of day, suggesting the localised presence of school based demand may impact store sales at particular times of the day. Of course this could be further correlated to school based demand through site surveys. However this is beyond the scope of this research. This micro-geographic change is also apparent across the entire study area and so temporal impacts upon stores will not occur just at a store level. The impacts are likely to be observed regionally too, which may explain why similar patterns are observed in all the cluster revenue profiles. For instance, the total night time population in West Yorkshire is 2.2 million and the daytime total is 2.4 million, with a net increase of over 200,000 individuals. While overall, West Yorkshire experiences a net increase in demand during the day, it is important to note that not all buffer zones gain in numbers. Several buffer zones (for different cluster groups) demonstrate a decrease in the total population during the day and this is seen most notably in cluster four's 0-0.5km buffer zone. Here the night time population is 292% of the daytime population, resulting from the departure of students out of the area. To reiterate, this may explain the observed temporal revenue profile of cluster four stores, which experience the majority of their sales in the evening when the population is nearby. This is in addition to the distinctive lifestyles and nightlife that students exhibit. Decreases may also account for periods of reduced revenue during the day as the spatiotemporal make up of catchment demand changes between phases of activity. The evidence presented across Tables 5.4 and 5.5, suggests why revenue predictions using counts of night time populations are insufficient and an inaccurate representation of daytime customer demand and how achievement of the thesis goals will likely result in not only a novel application of data and improved modelling, but possibly the initiation of positive change in industry practise.

Table 5.4 - Count of demand types overnight within buffer zones by each cluster group of stores

Buffer distance (km)	Night time residential demand								Total residential count
	Resident count: Non-student/unemployed or retired	Presence in buffer (% of total)	University demand count	Presence in buffer (% of total)	School based demand count	Presence in buffer (% of total)	Employed resident count	Presence in buffer (% of total)	
	Cluster group 1								
0 - 0.5	10894	32	9343	27	2374	7	11785	34	34396
0.5 - 1	49320	44	14181	13	9427	8	39879	35	112807
1 - 2	156849	47	36024	11	31127	9	108504	33	332504
2 - 5[1]	357844	51	21824	3	70739	10	248210	36	698617
5 -10	266721	50	14149	3	49732	9	202358	38	532960
10 - 15	173512	53	13070	4	30257	9	112144	34	328983
15 - 20	67191	55	2221	2	11314	9	41127	34	121853
20 - 45	37260	58	1176	2	6081	10	19421	30	63938
Subtotal[2]	-	**49**[3]	-	**7**	-	**10**	-	**35**	-
	Cluster group 2								
0 - 0.5	14144	55	777	3	2402	9	8521	33	25844
0.5 - 1	38071	50	5187	7	6437	8	26631	35	76326
1 - 2	140390	50	11993	4	27535	10	101461	36	281379
2 - 5[1]	553234	48	70871	6	106313	9	412009	36	1142427
5 -10	269929	52	19853	4	51955	10	179506	34	521243
10 - 15	75189	57	2406	2	12212	9	41941	32	131748
15 - 20	28634	61	901	2	4197	9	13359	28	47091
20 - 45	0	0	0	0	0	0	0	0	0
Subtotal[2]	-	**49**[3]	-	**6**	-	**9**	-	**36**	-

Table 5.4 - continued

	Cluster group 3								
0 - 0.5	29678	50	3893	7	4993	8	20423	35	58987
0.5 - 1	49768	52	3750	4	8792	9	34318	36	96628
1 - 2[1]	117133	44	30521	12	21680	8	94626	36	263960
2 - 5	368756	50	36603	5	67338	9	269439	36	742136
5 -10	370269	51	23061	3	72558	10	253557	35	719445
10 - 15	165060	53	13427	4	32565	10	101439	32	312491
15 - 20	18927	58	733	2	3125	10	9626	30	32411
20 - 45	0	0	0	0	0	0	0	0	0
Subtotal[2]	-	**47**[3]	-	**9**	-	**8**	-	**36**	-
	Cluster group 4								
0 - 0.5	381	2	12548	66	346	2	5815	30	19090
0.5 – 1[†]	5070	21	10254	42	875	4	7957	33	24156
1 - 2	15575	34	11164	25	2827	6	15828	35	45394
2 - 5	127468	48	16342	6	22941	9	96305	37	263056
5 -10	177176	50	9240	3	34269	10	130523	37	351208
10 - 15	270962	51	21488	4	54194	10	180177	34	526821
15 - 20	207372	52	9524	2	38936	10	144590	36	400422
20 - 45	315587	53	21428	4	56663	10	202233	34	595911
Subtotal[2]	-	**13**	-	**53**[3]	-	**3**	-	**32**	-

Table 5.5 - Count of demand types during the day within buffer zones by each cluster group of stores

	Total residential count		Residential demand		University demand				School based demand		Workplace demand		Leisure demand		
Buffer distance (km)	Traditional residential count	Presence in buffer (% of total)	Daytime demand	Presence in buffer (% of total)	At residence[4]	Presence in buffer (% of total)	On Campus[4]	Presence in buffer (% of total)	At school	Presence in buffer (% of total)	At Workplace	Presence in buffer (% of total)	At Attractions	Presence in buffer (% of total)	Total day time count
	Cluster group 1														
0 - 0.5	34396	30	10894	10	934	1	10368	9	900	1	76071	67	13898	12	113065
0.5 - 1	112807	75	49320	33	1418	1	0	0	10816	7	84304	56	4202	3	150060
1 - 2	332504	82	156849	39	3602	1	52403	13	22049	5	160758	40	10876	3	406537
2 - 5 [1]	698617	98	357844	50	2182	0	0	0	47279	7	295656	41	11599	2	714560
5 -10	532960	105	266721	52	1415	0	2302	0	38427	8	196617	39	2558	1	508040
10 - 15	328983	84	173512	44	1307	0	17658	5	19348	5	159849	41	19573	5	391247
15 - 20	121853	102	67191	56	222	0	0	0	7558	6	42705	36	2231	2	119907
20 - 45	63938	103	37260	60	118	0	0	0	4635	7	20098	32	0	0	62111
Subtotal[2]	-	**85**	-	**42**	-	**1**	-	**5**	-	**6**	-	**45**[3]	-	**3**	-
	Cluster group 2														
0 - 0.5	25844	39	14144	21	78	0	0	0	2456	4	49675	75	0	0	66353
0.5 - 1	76326	52	38071	26	519	0	17658	12	5651	4	81263	55	3557	2	146719
1 - 2	281379	111	140390	55	1199	0	0	0	15831	6	91156	36	5573	2	254149
2 - 5 [1]	1142427	92	553234	45	7087	1	55557	4	77703	6	527307	42	20470	2	1241358
5 -10	521243	91	269929	47	1985	0	10368	2	39386	7	217935	38	35337	6	574940
10 - 15	131748	102	75189	58	241	0	0	0	7286	6	46711	36	0	0	129427
15 - 20	47091	88	28634	54	90	0	0	0	2699	5	22011	41	0	0	53434
20 - 45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Subtotal[2]	-	**89**	-	**44**[3]	-	**1**	-	**4**	-	**6**	-	**44**[3]	-	**2**	-

Table 5.5 - continued (WPZ appear to have a considerable impact so we give this further attention in chapter seven)

	Cluster group 3														
0 - 0.5	58987	66	29678	33	389	0	0	0	4141	5	49457	56	5392	6	89057
0.5 - 1	96628	66	49768	34	375	0	0	0	5452	4	90832	62	0	0	146427
1 - 2 [1]	263960	87	117133	39	3052	1	28980	10	11103	4	136177	45	6433	2	302878
2 - 5	742136	91	368756	45	3660	0	44235	5	52928	6	329689	40	20721	3	819989
5 -10	719445	97	370269	50	2306	0	0	0	52131	7	296912	40	23885	3	745503
10 - 15	312491	96	165060	51	1343	0	10368	3	23674	7	117651	36	8506	3	326602
15 - 20	32411	90	18927	53	73	0	0	0	1583	4	15340	43	0	0	35923
20 - 45	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Subtotal[2]	-	**78**	-	**37**	-	**1**	-	**5**	-	**4**	-	**51**[3]	-	**2**	-
	Cluster group 4														
0 - 0.5	19090	292	381	6	1255	19	0	0	0	0	4911	75	0	0	6547
0.5 – 1†	24156	60	5070	13	1025	3	27927	69	387	1	5838	15	0	0	40247
1 - 2	45394	43	15575	15	1116	1	24476	23	1017	1	59640	56	3973	4	105797
2 - 5	263056	81	127468	39	1634	1	0	0	13417	4	170796	53	9594	3	322909
5 -10	351208	95	177176	48	924	0	3155	1	24297	7	152807	41	10271	3	368630
10 - 15	526821	90	270962	46	2149	0	10368	2	39956	7	246386	42	14894	3	584715
15 - 20	400422	99	207372	51	952	0	0	0	37072	9	151206	38	6115	2	402717
20 - 45	595911	94	315587	50	2143	0	17658	3	34866	5	244474	39	20090	3	634818
Subtotal[2]	-	**92**	-	**12**	-	**5**	-	**60**[3]	-	**1**	-	**23**	-	**0**	-

[1] Core catchment defined by distance that a total of 70% of sales is achieved.

[2] Proportion of demand type within core catchment (% of total).

[3] Main demand group based on total volume in core catchment (% of total).

[4] Estimates of university demand are based on a survey identifying the proportion of students on campus throughout the day. Values are based on midday estimates reaching around 90% of the total (Charles-Edwards and Bell, 2013).

† Speculative core catchment estimated from the spatial concentration of loyalty card sales and customers around stores within West Yorkshire (shown in Figures 4.7 and 4.8).

5.5 Conclusions

This chapter has built on the arguments and analysis of chapter four which looked at temporal revenue fluctuations by providing evidence of fluctuating demand which can be observed to have an impact upon store revenue. The aim of the chapter was to clearly identify different demand type geographies, with different behavioural characteristics and spatiotemporal patterns throughout the day, which are likely to have temporal impacts upon store revenue. The existence of a temporal demand was first indicated using observed consumer data where loyalty card transactions demonstrated that in addition to habitual behaviour, the interaction of consumers with grocery stores in some instances appeared not to be linked to a local residential store or to follow the 'traditional big weekly' shop format. It should be noted that in some regard loyalty card data provide unprecedented access and detail on consumer behaviour and consumption patterns (Demoulin and Zidda, 2008, Mauri, 2003), and moreover the fact that access to commercially sensitive datasets of this type are restricted in academic research making this research novel, loyalty card data is not without limitations. Lack of market representation, (i.e. detailing on only one retailer) is a problem (Birkin et al., 2010), although the additional detail and insight outweighs the potential associated bias.

The effects of spatiotemporal fluctuations regarding consumer demand were examined further, looking at specific demand types and at their assumed spatiotemporal movements using a range of open source and national datasets. The analysis of this data revealed considerable temporal components to each of the demand types, resulting in extensive redistribution of populations throughout the day, demonstrated across a day and night cycle. In addition to demonstrating a considerable spatiotemporal component to consumer demand, the evidence presented in this chapter amplifies the agenda of this research, building on the lack of representation and understanding of spatiotemporal consumer demand, particularly in location planning. It is likely that limited representation will not only result in poorer predictive accuracy due to the oversimplification of demand behaviour, but also quite clearly does not begin to account for the high degree of temporal change that takes place during the day as demand shifts. For instance, there is a clear distinction between the location and distribution of a daytime and night time population. During the day demand types were observed to move towards more urban areas such as large towns and city centres and to become highly concentrated in localised areas. This tendency was observed in multiple demand types, albeit for different reasons, resulting in diverse and unique distributions of consumers e.g. clustering at a place of work, around a tourist attraction or at a school. Moreover, each demand type's unique spatial distribution is likely to follow individual temporal patterns according to their own time constraints e.g. workers constrained by contracted hours or school students by school hours. The spatiotemporal distribution of various demand types, exhibited on a diurnal scale, are presented below in Figures 5.19 and 5.20. The maps (using the data presented in chapter seven) are representative of the assumed spatiotemporal distributions of each demand type using daytime and night time periods for the Leeds LAD. Understanding these patterns and being able to model any variations that occur will help to improve store level insight and analyses.

In the next series of chapters, I will build a disaggregated SIM of West Yorkshire, adding in temporal components of demand with each incremental and novel extension of the SIM. Within the scope of this research, this approach will demonstrate the shortcomings of traditional demand layers, as well as demonstrating the improvements in accuracy, presenting an interesting opportunity for insight into the impact of time on grocery store revenue.

Figure 5.19 - Midday (‘day time’) demand distribution estimates

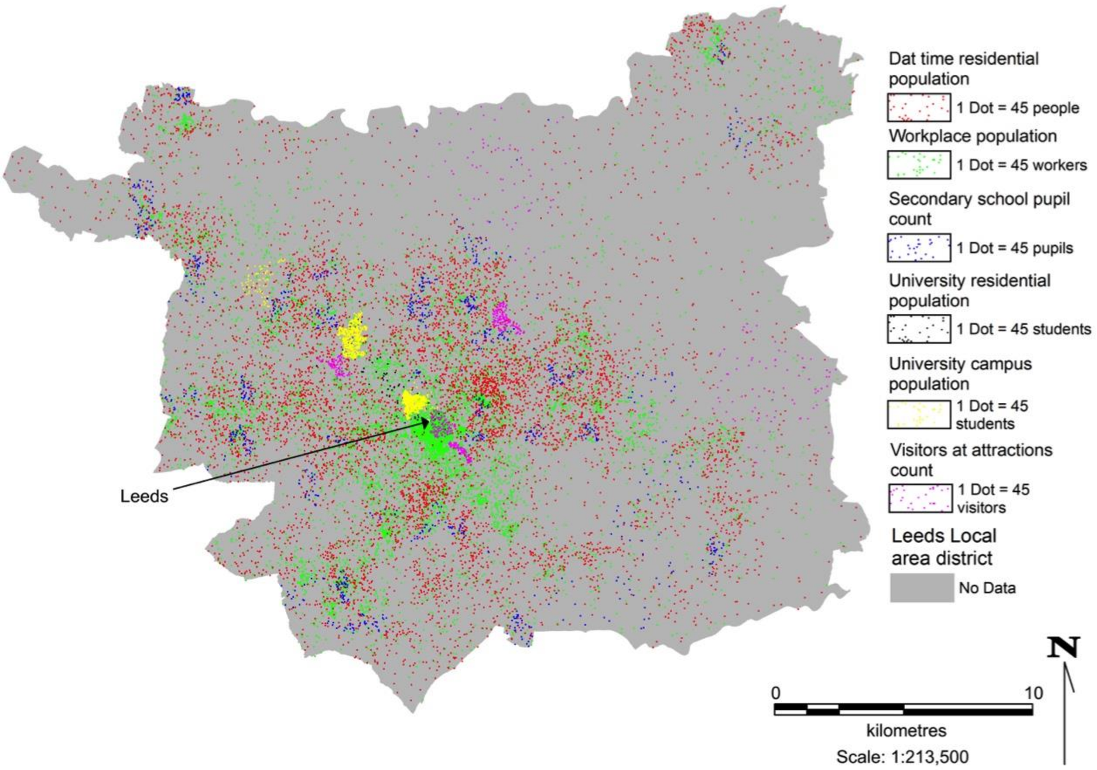
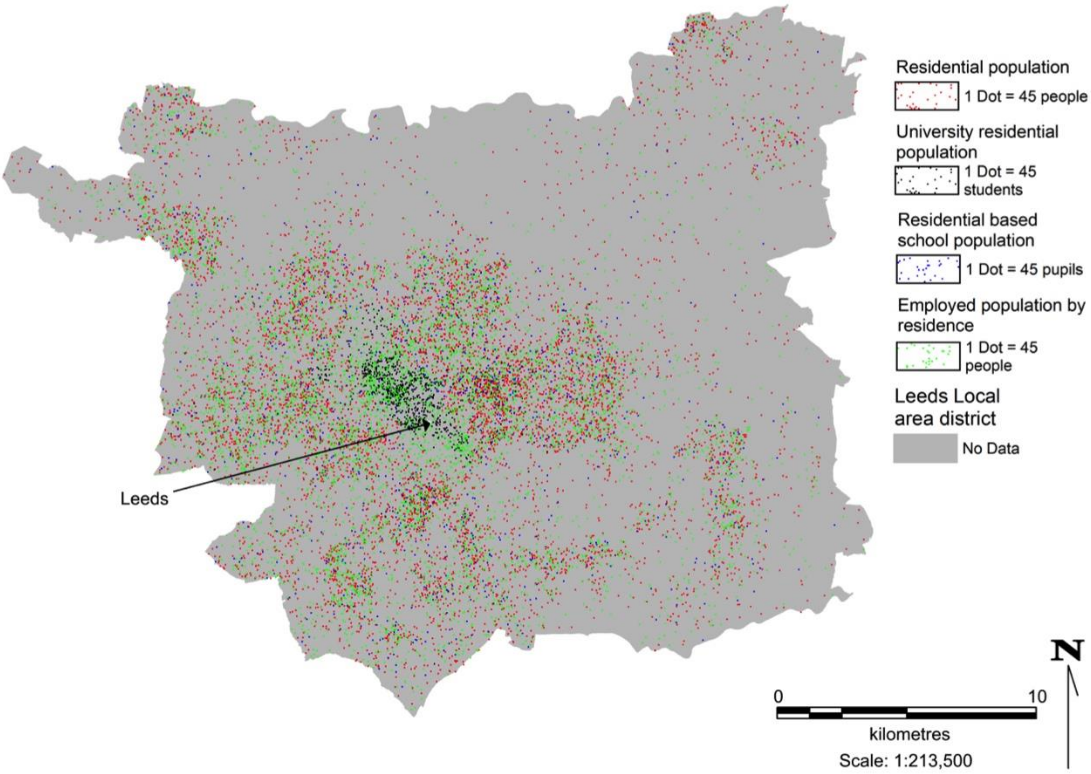


Figure 5.20 - Night time demand distribution estimates



Chapter 6 – A Disaggregated Spatial Interaction Model (SIM): For the grocery industry

6.1 Introduction

The role of SIMs and their importance in location planning has been identified in Chapter three. The discussion noted the wide uptake of the SIM as well as highlighting areas of current weakness, some of which were identified as potential opportunities for development and refinement. To recap, these areas of weakness in the conventional model relate to a poor capacity to represent more complex and diverse aspects of spatiotemporal consumer behaviour, and the impacts for store revenues. The use of the production-constrained SIM for the modelling of consumer expenditure flows particularly by location analysis teams in the grocery industry was specifically highlighted. The second half of chapter three introduced the aggregate production-constrained model in its conventional form as well as a discussion of approaches for modelling temporal and complex consumer flows. The next few chapters seek to demonstrate an improved predictive capacity of the SIM, by incorporating additional consumer types and behaviours through new and current supply and demand side parameters. The benefits that disaggregation can provide to a model have previously been noted in Chapters two and three and it is known to be an important stage for accurately reproducing observed consumer behaviour and for predicting store revenue (Benoit and Clarke, 1997, Birkin et al., 2010, Newing et al., 2014b). The process of SIM disaggregation will be discussed in the following chapter.

This chapter will focus on building a detailed SIM through the initial development of an aggregate, more traditional model. The modelling of the temporal impacts on store performance and the benefits of such modelling will be introduced in Chapter seven. This will be achieved firstly through extensive disaggregation of the model and then by the incorporation of a detailed additional workplace demand layer demonstrating the enhancements to store predictions through the incorporation of temporal behaviour, and distribution, associated with their different consumer types, other temporal demand layers will be added in chapter 8.

6.2 Disaggregation of the model

SIMs can be disaggregated to account for more complex consumer behaviour and it is realistic to assume that consumers will demonstrate different behaviour according to a range of factors such as diverse consumer characteristics, as discussed in Chapter two. It is these differences that influence consumers' shopping habits. It is widely reported that geodemographic characteristics of consumers impact consumer behaviour, influencing service, brand and product choice. Consequently, the propensity to travel and store choice are dependent on, and will vary according to, a series of socio-economic characteristics (Kamarulzaman, 2010, Newing et al., 2014b, Thompson et al., 2012). Therefore, the segmentation of consumers by geodemographic status in a SIM is an important component for disaggregation and calibration (Kamarulzaman, 2010). It is also reasonable to assume that once disaggregated by consumer type it is possible to replicate individual consumer behaviour through changes to the supply and demand parameters α and β, which control these within the model. For example, research has shown that consumer types will find different supermarket brands attractive and that certain types of consumer will be more attracted to a particular brand (Benoit and Clarke, 1997, Thompson et al., 2012). Similarly, the assumption that different consumer types will travel more or less is also widely accepted. Therefore, disaggregation by consumer type and (calibration of these parameters) will likely boost the potential for SIMs to relate consumer behaviour better. Disaggregation has been separated into two parts: first is the disaggregation and calibration of beta in sections 6.2.1 and 6.2.2 respectively, which is then followed by the disaggregation and calibration of alpha in sections 6.2.3 and 6.2.4.

The aggregate production-constrained SIM can be expressed below as:

$$S_{ij} = A_i \, O_i \, W_j^{\alpha} exp^{-\beta c_{ij}} \tag{6.1}$$

S_{ij} is the flow of expenditure between origin i and store j.

O_i represents the demand or amount of expenditure available in origin i: this is derived from the segmentation of average weekly spending by output area classifications at household level, generated in basic form by equation (6.2).

$$O_i = \sum_g H_i^g F^g \tag{6.2}$$

where,

H_i^g is the total number of households in origin i by household classification type g and

F^g is the average weekly spend on groceries by the household classification type g.

Expenditure estimates are derived from the average weekly household expenditure by OAC supergroup (using the most recent classifications and survey results) obtained via the Living Costs and Food Survey 2015 (ONS, 2015a).

W_j^{α} is the measure of attractiveness of store j (e.g. store size), raised to the power of α, where α is a parameter influencing the importance of the attractiveness W for store j.

$exp^{-\beta C_{ij}}$ is the distance deterrence term indicating the propensity to travel; β is the distance decay parameter and controls flows by influencing the importance of distance. C_{ij} accounts for the cost, the distance or travel time between origin i and store j.

A_i is a balancing factor and ensures that all demand in the area is allocated to centres within the model, and is shown in equation (6.3).

$$A_i = \frac{1}{\sum_j W_j \times exp^{-\beta C_{ij}}} \tag{6.3}$$

A refinement to the estimation of expenditure O_i to include a proportion of weekly alcohol spend was also incorporated at this stage to improve expenditure estimates. This was made under the assumption that some portion of weekly alcohol spending is spent at supermarkets. According to a report published by the Scottish Government this makes up 80% of off-trade sales (Scottish Government, 2007) with major supermarkets being the dominant supplier in off-trade sales and was shown to continue to increase (Alcohol Concern, 2011, Euromonitor International, 2015). While evidence clearly identifies supermarkets as the dominant market for off-trade sales, the actual proportion of on-trade and off-trade sales is referred to as an equal split with several press releases presenting corresponding evidence (Alcohol Concern, 2011, Ginley, 2013). It should be noted that this is likely to change as current trends suggest an increasing off-trade market with consumers buying higher volumes of alcohol and increasingly drinking at home (Mintel, 2010). Therefore, the relative value for weekly alcoholic drink spend for each household classification was also identified in the ONS 'Living Costs and Food Survey' (ONS, 2015a) (formerly the 'Expenditure and Food Survey'). Half of the available value was then added to the 'Food and non-alcoholic drinks' weekly spend by each output area classification type to generate an estimate of the total weekly expenditure in the grocery sector, as shown in equation 6.4.

The refinement of O_i is written as:

$$O_i = \sum_g H_i^g \left(F^g + \left(\frac{A^g}{2} \right) \right) \tag{6.4}$$

where in this instance; A^g represents the total weekly average alcohol spending for the output area classification type g, for each household.

6.2.1 Disaggregation by household type using 2011 OAC

The segmentation of customers was achieved via demand and geodemographic data which are nationally recognised sources of geographic data; the 'Output area classification' (OAC) utilised by the Office for National Statistics (ONS) and the 2011 census. The OAC is a unique tool providing insight into the population identifying similar and consistent characteristics between regions that are used to classify households into groups and can be used in retailing to identify consumer behaviour (Kamarulzaman, 2010, Vickers and Rees, 2006). In a previous release, the OAC was the classification of all 175,434 OAs in England and Wales with an average of 124 households per output area, with households classified according to 41 variables derived from 2001 census data (Vickers and Rees, 2006). However, since its initial release in 2005, an updated version has been released via the ONS, based on 2011 Census data with updates to OA geographies and OAC 'supergroup' classifications for households (ONS, 2011a). In the updated release eight OAC supergroups (instead of the original seven) are now used for household classification. The use of OACs have been shown to be a sufficiently detailed segmentation of consumer types and it is widely accepted that OACs are a realistic variable for estimating the distance deterrence parameter β when disaggregating consumer behaviour in a SIM (Newing et al., 2013a, Thompson et al., 2012). Equally, OAC is an accredited national statistic based entirely on census data, providing robust household level socioeconomic and geographical coverage. Consequently, the supergroup classification of households will be adopted as the level of detail for SIM disaggregation within this research. A full breakdown of the variables and the weightings for household characteristics that led to the allocation of OAC supergroups to household and the current UK coverage for 2011 can be accessed on the ONS website (ONS, 2011a).

The disaggregation by consumer type in the model can be written as:

$$S_{ij}^{g} = A_i^{g} \, O_i^{g} W_j^{\alpha} exp^{(-\beta^{g} c_{ij})} \tag{6.5}$$

where variables are as before, but now disaggregated by household type g as shown below:

$exp^{(-\beta^{g} c_{ij})}$ a distance deterrence factor has been adapted, incorporating changes in behaviour for household type g, impacting the distance travelled between origin i and retail destination j.

S_{ij}^{g} represents predicted expenditure between origin i and store j by household type g.

O_i^{g} represents the amount of expenditure available in origin i by household type g.

W_j^{α} is the measure of the attractiveness of store j raised to the power of α (further disaggregation by brand, is discussed later).

A_i^g represents the balancing factor controlling competition in the model ensuring that all demand from origin i by household classification g is distributed to stores within the study area:

A_i^g, can be rewritten as follows:

$$A_i^g = \frac{1}{\sum_j W_j^{\alpha} exp^{(-\beta^g c_{ij})}} \tag{6.6}$$

6.2.2 Beta Calibration, Goodness- of-Fit results and model output

Consumers were disaggregated according to the 2011 OAC supergroups at LSOA geography and calibration was then undertaken to find appropriate β^g values. It should be noted that the calibration of SIMs using observed consumer data from a commercial source represents a unique opportunity for SIM research. The use of such commercial data has been previously limited. Its availability helps to produce novel developments as well as considerably improve the ability to reproduce observed consumer behaviour through the calibration of known consumer flow data and observed store revenues. Retail applications of SIMs, identifying both commercial retail applications and academic development of SIMs have highlighted this current lack of research utilising commercial data for calibration, which can lead to poorer model fits (Birkin et al., 2010, Guy, 1991).

For this research observed data have been provided by a major UK grocery retailer. Observed demand data consists of consumer loyalty card transaction records and weekly sales profiles for the West Yorkshire region. This data was used to establish observed flows and revenue figures vital in the calibration process. Observed consumer flow data were extracted from the extensive loyalty card dataset, (consisting of approximately 29 million individual records) using the programming language R, (to LSOA and store level for the study area). Due to the nature of the data, observed flows are only generated by loyalty card customers and therefore all other non-loyalty card consumer spending is not accounted for. Using observed revenue data, the proportion of total revenue from loyalty card flows to stores was calculated. A calibration sample was chosen with the following requirement: stores with a minimum of 30% of weekly revenue being generated by loyalty card customers. Observed flows were then up scaled to be representative of all grocery spending (for this sample of stores) so that they could then be used for calibration. A smaller validation subset was taken from this data and excluded for the calibration process so that they could be used as a control group after calibration to ensure that actual observed consumer behaviour has been replicated and that the model has not been artificially fitted to the observed data.

Average trip distance (ATD) was calculated for both observed and predicted flows as previously demonstrated in Chapter three. The calibration of a disaggregated consumer model (using OACs) was calculated by applying incremental changes to beta values until the predicted behaviour correlated with the observed. The calibration of predicted ATD adopted a series of statistical methods as well as assessment of the spatial distribution of ATD using flow maps. This was to ensure an optimal beta value was identified, repeating the process for each OAC supergroup and this is expressed below (equation 6.5).

Observed and predicted ATD were first balanced against each other,

$$ATD^{O_g} = \frac{ATD^g}{\widetilde{ATD}^g} \tag{6.7}$$

where:

ATD^{O_g} represents the optimised ATD by household type g, where the difference between predicted ATD and observed ATD is minimised. This is balanced through the estimation of β by OAC household type g until ATD^g and $\widetilde{ATD}^g$ are as similar as possible. A value of 1 represents an exact match.

and where,

ATD^g represents predicted ATD by OAC household type g and $\widetilde{ATD}^g$ represents observed ATD by OAC household type g (derived the using loyalty card records of actual customers), respectively calculated as follows:

$$ATD^g = \sum_b \frac{\sum_{ij} S_{ij}^{gb} C_{ij}}{\sum_{ij} S_{ij}^{gb}} \tag{6.8}$$

$$\widetilde{ATD}^g = \sum_b \frac{\sum_{ij} \check{S}_{ij}^{gb} C_{ij}}{\sum_{ij} \check{S}_{ij}^{gb}} \tag{6.9}$$

in which; S_{ij}^{gb} represents predicted expenditure flows and $\check{S}_{ij}^{gb}$ which represents observed expenditure flows, with b representing brand type.

The sum of squared errors was also used to balance observed and predicted ATD and is expressed below.

$$SSE = \sum_{ig} \left(ATD_i^g - \widetilde{ATD_i^g} \right)^2$$

(6.10)

Correlation was used to assess the strength of the relationship between observed and predicted ATD and is expressed as:

$$Corr\left(ATD^g, \widetilde{ATD^g}\right) = \frac{COV\left(ATD^g, \widetilde{ATD^g}\right)}{\sigma(ATD^g)\sigma(\widetilde{ATD^g})}$$

(6.11)

where σ is the standard deviation.

Finally, R^2 expressed in equation 3.10 was also used for calibration. The combination of statistical methods and GIS visualisation in the calibration process was used to ensure a high level of robustness when estimating beta values. Beta values were incrementally changed until an ‘optimal’ beta value for each OAC was identified. Table 6.1 below demonstrates that a beta value of 0.43 for the aggregate population was optimal. However as expected, when disaggregated by consumer type, ATD^{O_g} are no longer balanced using a single aggregate beta. Table 6.2 shows examples of disaggregated beta values for each OAC. Each OAC was again calibrated until the ATD^{O_g} value was as close to one as possible, signifying a correlating value, between observed and predicted ATD. The output of optimised values for beta were then plotted via GIS. Figures 6.1 and 6.2 show the spatial distribution of ATD with comparable spatial patterns for observed and predicted ATD, strengthening this methodology. This indicates that the model is capable of producing realistic replications of consumer behaviour regarding distance travelled to grocery stores. Following this stage of calibration, the optimal beta values were then used in the model. The calibration of ATD, using observed loyalty card data, for each origin means that it was also possible to account for the differences is distance travelled to a store for rural and urban customers. Regardless of OAC rural customers are likely to travel further to do grocery shopping because there are typically fewer options available. However, by using observed consumer shopping patterns to calibrate ATD in the model it is possible to incorporate these differences in consumer behaviour.

Table 6.1 - Observed and predicted ATD for West Yorkshire LSOAs by OAC using an aggregate beta value for the entire population.

Aggregate consumer			
Straight line distance (km)			
An optimal beta of 0.43 produced a ATD^{O_g} of 1.0			
OAC	ATD^{g}	$ATD^{\check{g}}$	ATD^{O_g}
1 - Rural residents	7.32	8.56	0.85
2 - Cosmopolitans	4.18	3.78	1.11
3 - Ethnicity central	3.94	2.94	1.34
4 - Multicultural metropolitans	3.12	2.03	1.53
5 - Urbanites	3.38	2.96	1.14
6 - Suburbanites	4.30	4.05	1.06
7 - Constrained city dwellers	4.19	3.79	1.11
8 - Hard-pressed living	4.82	4.27	1.13

Table 6.2 - Observed and predicted ATD for West Yorkshire LSOAs using disaggregated beta values by OACs.

Disaggregated consumer types				
Straight line distance (km)				
OAC	Beta examples	ATD^{g}	$ATD^{\breve{g}}$	ATD^{O_g}
1 - Rural residents	0.32	8.56	8.57	1.00
2 - Cosmopolitans	0.53	3.78	3.79	1.00
3 - Ethnicity central	0.63	2.94	2.97	1.01
4 - Multicultural metropolitans	0.68	2.03	2.05	1.01
5 - Urbanites	0.5	2.96	2.97	1.00
6 - Suburbanites	0.46	4.05	4.08	1.01
7 - Constrained city dwellers	0.49	3.79	3.75	0.99
8 - Hard-pressed living	0.49	4.27	4.32	1.01

Figure 6.1 - (below) shows the observed ATD in West Yorkshire for consumers disaggregated by OAC.

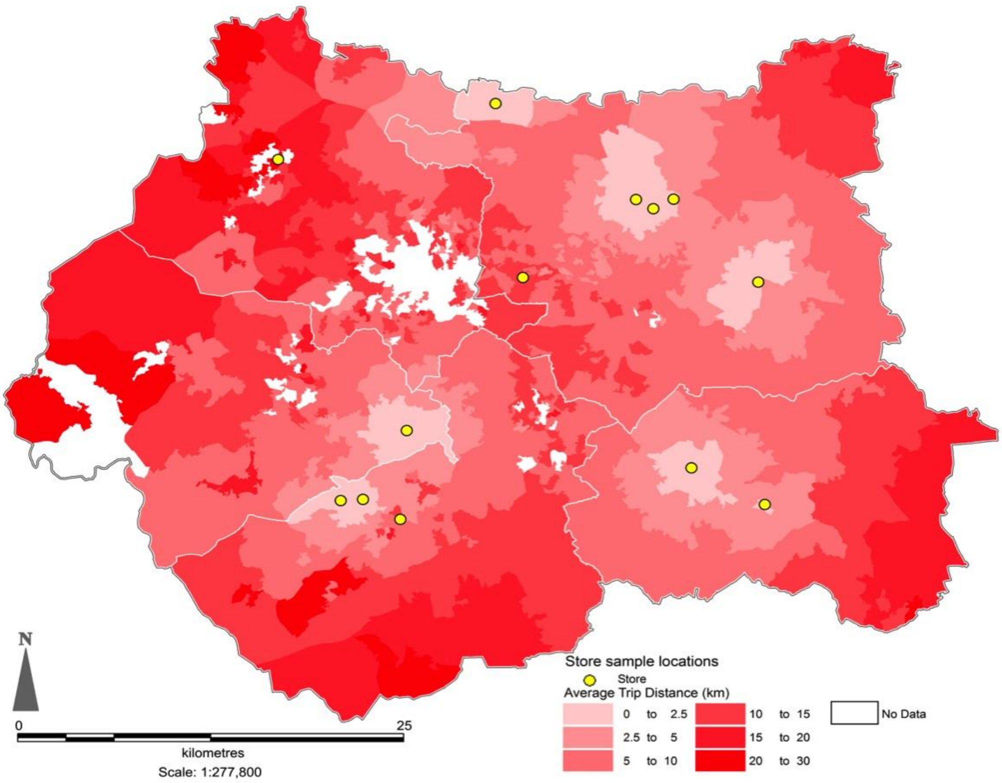
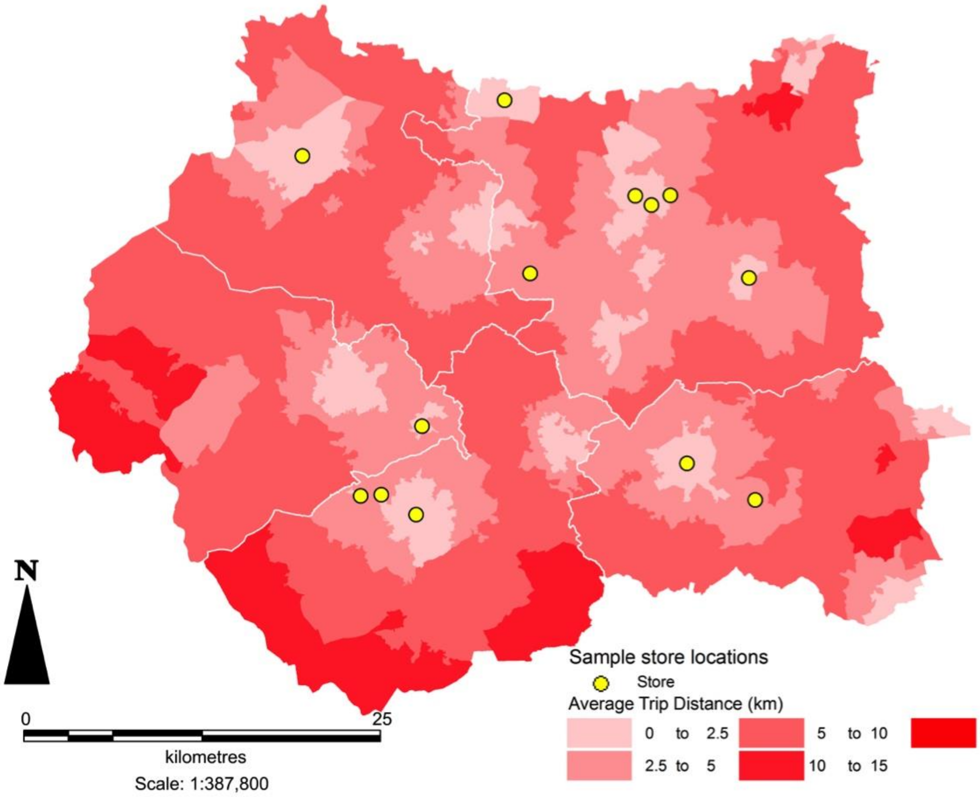


Figure 6.2 - (below) shows the predicted ATD (with optimised beta) in West Yorkshire for consumers disaggregated by OAC.



Following disaggregation of beta by consumer type, R^2 and SRMSE (expressed in equations 3.10 and 3.9, respectively) were used for an assessment of current model performance, using observed and predicted flows. The goodness-of-fit (GOF) statistics for this stage are shown in Table 6.4 below. The model overall seems to underperform, falling short of typical commercially expected thresholds (within 10% of actual revenue) used by retailers when predicting store revenue using a SIM. The scale of under prediction of flows to stores is shown below in Figure 6.3. While the GOF for flows following disaggregation by OAC shows improvement compared to the GOF for the aggregate model, (shown below in Table 6.3) the continued under-predictions suggests that there is a clear need for further refinements to the model. This is to ensure an improved level of accuracy if it is to be used for real world applications. Thus, the next stage involved the disaggregation of alpha (store attractiveness) by brand type (shown below). It has been demonstrated that consumer types will favour different grocery retailers (Newing et al., 2014b, Thompson et al., 2012). By recreating observed consumer brand preference it is assumed that the distribution of flows to retailers will be more accurate and improve model performance (discussed in section 6.2.3 and 6.2.4). While disaggregation of alpha is likely to improve the accuracy of the model, it is unlikely that this will fully resolve the current level of underperformance in most store revenue predictions (and the smaller number of over-predictions), which may be the result of spatiotemporal demand inaccuracies. In this instance, if proven that the model is recreating observed consumer behaviour, then it is reasonable to assume that poor performance is due to a poor temporal representation of demand in the model, which is summarised in section 6.3.

Figure 6.3 – Observed and predicted store revenue for the sample of stores in the aggregate and disaggregate SIMs.

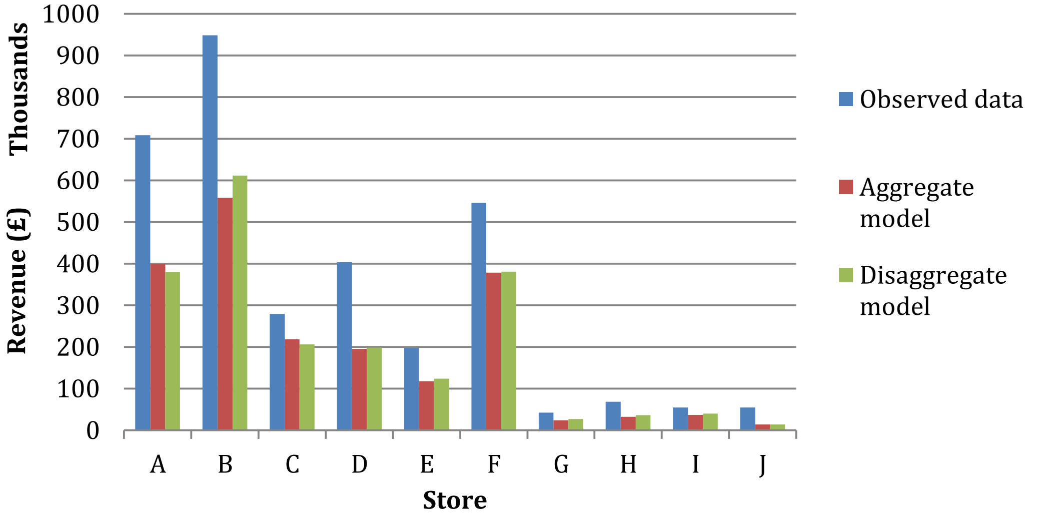


Table 6.3 - GOF statistics on observed and predicted flows (£) for aggregate model.

Aggregated consumer		
GOF assessment on flows (£)		
Aggregate beta set at 0.43		
OAC	R^2	SRMSE
1 - Rural residents	0.79	1.75
2 - Cosmopolitans	0.48	5.03
3 - Ethnicity central	0.89	7.59
4 - Multicultural metropolitans	0.78	4.26
5 - Urbanites	0.89	2.83
6 - Suburbanites	0.86	3.00
7 - Constrained city dwellers	0.74	3.91
8 - Hard-pressed living	0.80	3.45
Overall GOF	0.82	3.55

Table 6.4 - GOF statistics on observed and predicted flows (£) for disaggregate model.

Disaggregated consumer			
GOF assessment on flows (£)			
OAC	Beta**	R^2	SRMSE
1 - Rural residents	0.32	0.79	1.85
2 - Cosmopolitans	0.53	0.58**	4.51
3 - Ethnicity central	0.63	0.96	8.52
4 - Multicultural metropolitans	0.68	0.91	2.54
5 - Urbanites	0.5	0.92	2.38
6 - Suburbanites	0.46	0.87	2.82
7 - Constrained city dwellers	0.49	0.76	3.73
8 - Hard-pressed living	0.49	0.83	3.12
Overall GOF		0.84	3.05

** – Calibration of the model parameters will be a continuous process during SIM development to ensure that the most appropriate parameter values continue to be used. Cosmopolitans represent an example of poor model fit, which will be addressed through on going calibration.

6.2.3 Disaggregation of alpha by retailer brand and household classification

It is a realistic assumption that the attractiveness of different retailer brands will be influenced by various socioeconomic and demographic consumer traits. Therefore, in order to replicate actual consumer behaviour within SIMs accurately, it is important to address this by disaggregating alpha by consumer type and supermarket brand. At this stage, the segmentation of consumer household types has been derived from the Office of National Statistics ‘OAC’ system release developed in 2006 (following the same disaggregation process for consumer type adopted for beta in section 6.2.1). Alpha was also disaggregated by brand. Grocery retailers, according to the OAC of a consumer, received an alpha parameter specific to their brand for that consumer type. This initial segmentation of brand attractiveness using alpha has been adapted from existing research on disaggregation in Newing et al. (2014. p9) and Thompson et al. (2012) in Table 6.5 below. The foundation of the alpha values used in their research were derived from extensive analysis into the representation of OAC supergroups and brand preference with individual major grocery retailers in the UK. The research made use of extensive consumer data from the private sector provided by the research company Acxiom Ltd, which detailed household spending habits through an extensive consumer lifestyle survey of approximately 750,000 UK households. Using the research opinion poll provided by Acxiom, Thompson et al. (2012) demonstrated the preferences for individuals to different major grocery retailers based on who used them for their weekly shops. This was identified for each retailer to each household OAC. This was seen to provide a good indicator of brand choice by household type (Newing et al, 2014). Thompson et al. (2012) used location quotients for each brand and OAC supergroup in order to establish the level of over or underrepresentation of each OAC, within the brand's customer profile (in turn identifying brand preference). Newing et al. (2014) adapted the location quotients produced by Thompson et al. (2012) to develop a series of alpha values representative of brand attractiveness for each retailer and each OAC usable in SIMs. As with Newing et al. (2014), floorspace is currently being used as the measure of attractiveness within the SIM, with the alpha values operating as a power function, raising the attractiveness of a retailer within the SIM for different household types. Maintaining the same relative difference between the location quotients developed by Thompson et al. (2012), Newing et al. (2014) developed a set of values in order to replicate the actual observed consumer behaviour identified from Acxiom's consumer survey data. The alpha values presented by Newing et al. (2014), adapted below, represent the initial values used for alpha within this SIM, (which were coincidently first developed for consumer preference and brand attractiveness within the Yorkshire and Humber area) and represent a suitable starting point for consumer preference. It is likely that brand attractiveness has undergone a certain level of change within the UK grocery market, both nationally and regionally since 2014, and therefore further calibration of the alpha values is presented in section 6.2.4 to ensure that they accurately replicate current consumer behaviour.

Table 6.5 - Alpha values for UK grocery retailers according to retailer brand and OAC adapted from Newing et al. (2014)

Table 1. Brand location quotients used to set alpha values

Brand (retailer)	OAC supergroup						
	1	2	3	4	5	6	7
	Blue collar	City living	Countryside	Prospering suburbs	Constrained by circumstances	Typical traits	Multicultural
Aldi	0.9980	0.9970	1.0051	0.9987	1.0025	1.0005	0.9952
ASDA	1.0076	0.9912	0.9904	0.9970	1.0023	0.9992	1.0013
Co-Op	1.0020	0.9990	1.0157	0.9922	1.0008	1.0000	0.9894
Lidl	1.0015	0.9995	1.0066	0.9962	0.9957	0.9997	1.0091
M&S	0.9891	1.0381	0.9967	1.0066	0.9952	1.0051	1.0003
Morrisons	1.0005	0.9942	0.9997	0.9987	1.0020	1.0005	0.9990
Sainsbury's	0.9904	1.0121	1.0013	1.0088	0.9942	1.0028	0.9997
Tesco	0.9992	0.9987	1.0071	1.0010	0.9965	0.9990	0.9985
Waitrose	0.9811	1.1000	1.0061	1.0124	0.9843	1.0023	1.0068
Iceland	0.9997	0.9982	1.0058	0.9975	0.9991	1.0001	1.0021

Source: Newing et al. (2014) p9.

The incorporation of attractiveness disaggregated by brand and household type results in an extension of the model in which the balancing factor A_i, demand O_i and supply W_j^{α} are altered to include the influence of different types of consumer behaviour, caused by household types g. This refinement accounts for the variation of relative brand attractiveness by different consumer household types by disaggregating the power function alpha by both store brand b and household types g (which can be written as α^{gb}). The incorporation of these additional variables helps to account for the variations in, and impact of, different consumer behaviour concerning store choice within the model. The disaggregated model is written as:

$$S_{ij}^{gb} = A_i^g \, O_i^g W_j^{\alpha^{gb}} exp^{(-\beta^g c_{ij})}$$

(6.11)

where:

S_{ij}^{gb} represents predicted expenditure between origin i and store j by household classification type g and store brand b.

O_i^g represents the amount of expenditure available in origin i by household classification type g.

$W_j^{\alpha^{gb}}$ is the measure of attractiveness of store j and α^{gb} is a power function influencing the importance of the attractiveness variable for store j by household classification type g and store brand b.

A_i^g represents the balancing factor ensuring that all demand from origin i by household classification g is distributed to stores within the study area which is calculated through the extension shown in the following equation:

$$A_i^g = \sum_b \frac{1}{\sum_j W_j^{\alpha^{gb}} exp^{(-\beta^g c_{ij})}} \tag{6.12}$$

6.2.4 *Alpha Calibration, Goodness-Of-Fit results and model output*

It is reasonable to assume that brand attractiveness, while still dependant on regional characteristics, has undergone a change since the alpha values first developed by Newing et al. (2014) and Thompson et al. (2012). Present day data on brand preference, found in research opinion polls, is currently unavailable for this work. In an effort to update brand preferences and calibration of these older alpha values changes were made over three key stages. By using the alpha values originally developed by Newing et al. (2014) and Thompson et al. (2012), it is possible to incorporate observed preferences for different OACs while updating brand attractiveness to account for more recent changes in retailer market penetration. However, it was important to update the alpha values for the 2006 OAC supergroups to the newer 2011 classifications; this was also repeated on beta to account for changes in demand. The 2011 OACs represent a marked change in LSOAs classifications, with original classifications demonstrating no notable similarities to new supergroup classification in some locations. To overcome this issue, the decision was made to weight the 2011 OAC supergroup alpha values, according to the share of the previous OAC classification of each LSOAs within the study area, shown below in Table 6.6. Alpha was then weighted according to the percentage of each 2006 OAC supergroup found in the 2011 OAC supergroups for West Yorkshire LSOAs. Based on the distribution of the OACs within the West Yorkshire region, new alpha values were derived from those displayed in Table 6.5, taking the observed percentage of each original OAC alpha value, adding together each proportion share of the previous values resulting in new alphas for each 2011 OACs to ensure a fair representation of the OAC populations. The update to alpha appeared to improve model accuracy for supermarkets and convenience stores predictions, changing from 79% and 43%, to 87% and 48% of total observed revenue, respectively. Interestingly, the GOF assessment (R^2 and SRMSE) demonstrated very little improvement in terms of observed and predicted flows, suggesting the need for further demand refinements, such as the incorporation of spatiotemporal demand layers, redistributing expenditure around the model. At this stage the model is now accounting for the most recent classification of households in the study area. However, further refinements are necessary to account for evolving behaviours regarding brand preference to ensure a high level of applicability and representation of current market conditions.

Table 6.6 – Lookup table for 2006 (Vickers and Rees, 2006) to 2011 OACs for West Yorkshire LSOAs where each of the 2011 supergroup classifications is broken down by the proportion of their structure derived from the previous classification groups (% of total).

		2011 OAC Supergroups							
		1	2	3	4	5	6	7	8
2006 OAC Supergroups	1	0.00	0.00	0.00	9.17	3.26	10.03	26.36	29.82
	2	0.00	84.62	15.56	4.59	4.89	1.43	2.73	0.00
	3	26.67	0.00	0.00	0.00	5.98	10.32	0.91	3.31
	4	53.33	0.00	0.00	11.93	25.54	35.82	3.64	11.45
	5	0.00	3.85	13.33	10.70	4.35	1.72	39.09	16.27
	6	20.00	0.00	0.00	22.02	53.26	40.11	25.45	34.34
	7	0.00	11.54	71.11	41.59	2.72	0.57	1.82	4.82

OAC Supergroup descriptions

1 - Blue Collar Communities	1 – Rural residents
2 - City Living	2 – Cosmopolitans
3 - Countryside	3 – Ethnicity central
4 - Prospering Suburbs	4 – Multicultural metropolitans
5 - Constrained by Circumstances	5 – Urbanites
6 - Typical Traits	6 – Suburbanites
7 - Multicultural	7 – Constrained city dwellers
	8 – Hard-pressed living

The next stage involved deriving regional market shares for each grocery retailer. This made it possible to assess current model performance by taking into account spatial variations in grocery market shares that occur regionally throughout the UK (Burt and Sparks, 2003b, Thompson, 2013, Hughes et al., 2009). It is reasonable to assume that if the predicted regional market shares in the model match observed market shares, then the level of retailer attractiveness in the model is potentially representative of current consumer preferences regarding grocery shopping. Regional market shares were derived using observed national market shares obtained from recently released data (Kantar-Worldpanel, 2015), and national floor space provided in the partner's dataset for the same time period (supply side data on grocery retailers was provided for 2013). It should be noted that all 'symbols and independents' and 'other retailers' stores were grouped together to correspond with data outputs by Kantar Worldpanel for national market share to ensure a corresponding methodology. The total national floorspace for each retailer brand was divided by their observed national market share at the time. This established the market share value (% of market) per sq./ft. of floorspace for each grocery retailer. A sum of the total regional floorspace was then calculated. Using the market share value of a single sq./ft. estimated in the previous step, regional market shares based on regional floorspace were derived by multiplying market share per sq./ft. of floorspace with the sum of total regional floorspace thus providing an estimate of regional market share percentage. This method is expressed below and correlates with a similar approach by Hughes et al. (2009), which measured regional market shares for retailers in terms of their share of sales areas, strengthening the methodology. Although to date this has received little academic attention, Burt and Sparks (2003b), similarly note the relationship between the level of a brand's presence and market share, corroborating the theory behind this approach. For instance, Tesco's high national market share (and generally high regional market shares) are likely the result of an extensive national store network suggesting that brand penetration is also an important indicator further supporting this methodology (Hughes et al., 2009, Thompson, 2013). The regional market shares, shown below in table 6.7, offer a potential estimate of market shares, taking into account the national strength of a brand but also accounting for regional monopolies in areas where retailers similarly have strong brand loyalty. For example, this is seen for Asda and Morrison's who dominate in the Yorkshire and Humber area or, Tesco and Sainsbury's in the South of England. The work of Thomson (2013) offered a detailed examination of these spatial variations, using consumer data collated from a research opinion poll (as detailed above), this was used as a comparison for the regional market shares generated in this research using floorspace and demonstrated corresponding findings.

$$MSsq/ft^b = \frac{Nms^b}{Nfs^b} \tag{6.13}$$

$$NRms^b = MSsq/ft^b * Rfs^b \tag{6.14}$$

$$Rms^b = (NRms^b / \sum_{b=1}^{n} NRms^b) * 100 \tag{6.15}$$

where,

Nms^b represents national market share for brand b,

Nfs^b represents total national floor space for brand b,

$MSsq/ft^b$ is the market share value of one square foot of floor space for brand b,

Rfs^b represents the total regional floor space for brand b,

$NRms^b$ is the regional proportion of national market share according to regional floor space for brand b, where n is the number of brands under consideration.

Rms^b represents the regional market share for brand b, up scaled from $NRms^b$.

Table 6.7 – Estimated regional market shares for the main grocery retailers in West Yorkshire.

Estimated regional market shares for the main retailers in West Yorkshire, UK derived from observed National market shares.		
Brand - b	National Market share[1] (%) - Nms	Regional Market Share (%) - Rms
Aldi	3.8	3.37
Alfred Jones[2]	2.1	0.001
Asda	17.2	19.68
Booths[2]	2.7	0.48
Co-Op	6.4	6.57
Costcutter[2]	2.1	0.20
Farmfoods[2]	2.1	0.15
Iceland	1.9	1.24
Lidl	3.1	3.44
Londis[2]	2.1	0.21
Morrisons	11.2	22.53
Nisa[2]	2.1	0.18
Ocado[2]	2.7	0.28
One Stop[2]	2.1	0.14
Premier[2]	2.1	0.41
Spar[2]	2.1	0.13
Tesco	30	20.95
Waitrose	4.9	1.59
Sainsbury's	16.7	15.00

1 – National market shares were obtained for the same time period (2013) that supply side data provides to ensure corresponding market presence and were published online (Kantar-Worldpanel, 2015).

2 – Fall under Kantar Worldpanel categories: symbols/independents and other and are given their collective market share values accordingly. Total floorspace for all symbols/independents and all others were used grouping brands together to correspond with Kantar Worldpanel when establishing regional market shares.

For clarity, from this point onwards the estimated regional markets shares (Table 6.7) derived using national market share will hence be referred to as 'observed' regional market shares and those generated within the model will be referred to as 'predicted' regional market shares. Overall the predicted regional market shares appear to replicate observed conditions, although some variance to those presented in Thompson (2013) were identified. Taking into account recent changes in grocery retail behaviour, such as consumer patronage, brand presence and loyalty and brand switching (to the discounters, for example), this is not unexpected although the effectiveness of calibration is somewhat limited by data availability on both the demand side (such as shares are only a proxy) and on the supply side (potentially using incomplete data sources). Through the calibration of regional market shares, alpha values were adjusted until consumer behaviour was as similar as possible for observed and predicted values (Table 6.8). A high correlation and R^2 value for observed and predicted regional market shares suggests that the model is replicating observed consumer brand preferences satisfactorily. Table 6.9 (below) indicates updated alpha values for brand attractiveness according to the 2011 OACs. The assumption that observed market behaviour was reproduced reliably was tested through an analysis of the spatial distributions of both observed and predicted market shares and flows for the sample stores (shown below in Figure 6.4).

Within the estimated regional market shares (Table 6.7) several major retailers were observed to have a substantially different national market share when compared to their regional estimate. For example, Tesco and Waitrose both had a lower market share regionally than their national percentages whereas, Morrisons had a higher regional market share. These estimates correspond with both Burt and Sparks (2003b) and Thompson (2013) findings on the UK grocery sector regional market shares. Tesco and Waitrose have less market presence in the north and have a much larger presence in the south; hence both are far more powerful, in far greater numbers and have higher market shares (Waitrose to a lesser extent) in the south of the UK. Similarly, Morrisons has a more extensive store network in the north and correspondingly is a much stronger brand in the north, achieving higher market shares in this area of the UK when compared to their national market share. These regional estimates presented in this thesis use more recent data than the previous regional market share analysis, offering novel insight into grocery retailers regional performance in West Yorkshire. Insight into the long term changes of regional market shares for the major grocery retailers can be seen from 1991, 1995, 1999 and 2011 presenting a historical snapshots of regional brand strength for the UK over the last couple of decades (Burt and Sparks, 2003b, Thompson, 2013). When used in conjunction with the findings of this research these figures offer a picture of long-term trends. The origins of the north and south divide for retailers such as Tesco and Morrison's are partially the result of retailer heritage, with brands originating from an area and then reinforced by strong brand loyalty over time (Thompson, 2013). Importantly, the estimated regional market shares (Table 6.7) demonstrate similar market share behaviour to those shown in Burt and Sparks (2003b, p240) and Thompson (2013, p129-131), these corresponding values (Table 6.7) further strengthening the updated brand attractiveness for grocery retailers in West Yorkshire. This provides a reasonable level of reliance for the estimated market shares (as well as adding integrity to the method of estimating regional market shares from floorspace), which was used for calibration of the SIM. The existence of spatial variations in market shares demonstrates the importance of using regional market shares for model calibration. If consumer behaviour and grocery brands are calibrated using national market shares, SIM revenue predictions will likely be less accurate.

Figure 6.4 (shown below) Key - Observed and Predicted market shares (%) and Flows (£) in West Yorkshire, UK.

Map A – Observed market share (%) for calibration sample

Map B – Predicted market share (%) for calibration sample

Map C – Observed flows (£) for calibration sample

Map D – Predicted flows (£) for calibration sample

Figure 6.4 - Observed VS Predicted market shares (%) and Flows (£) in West Yorkshire, UK.

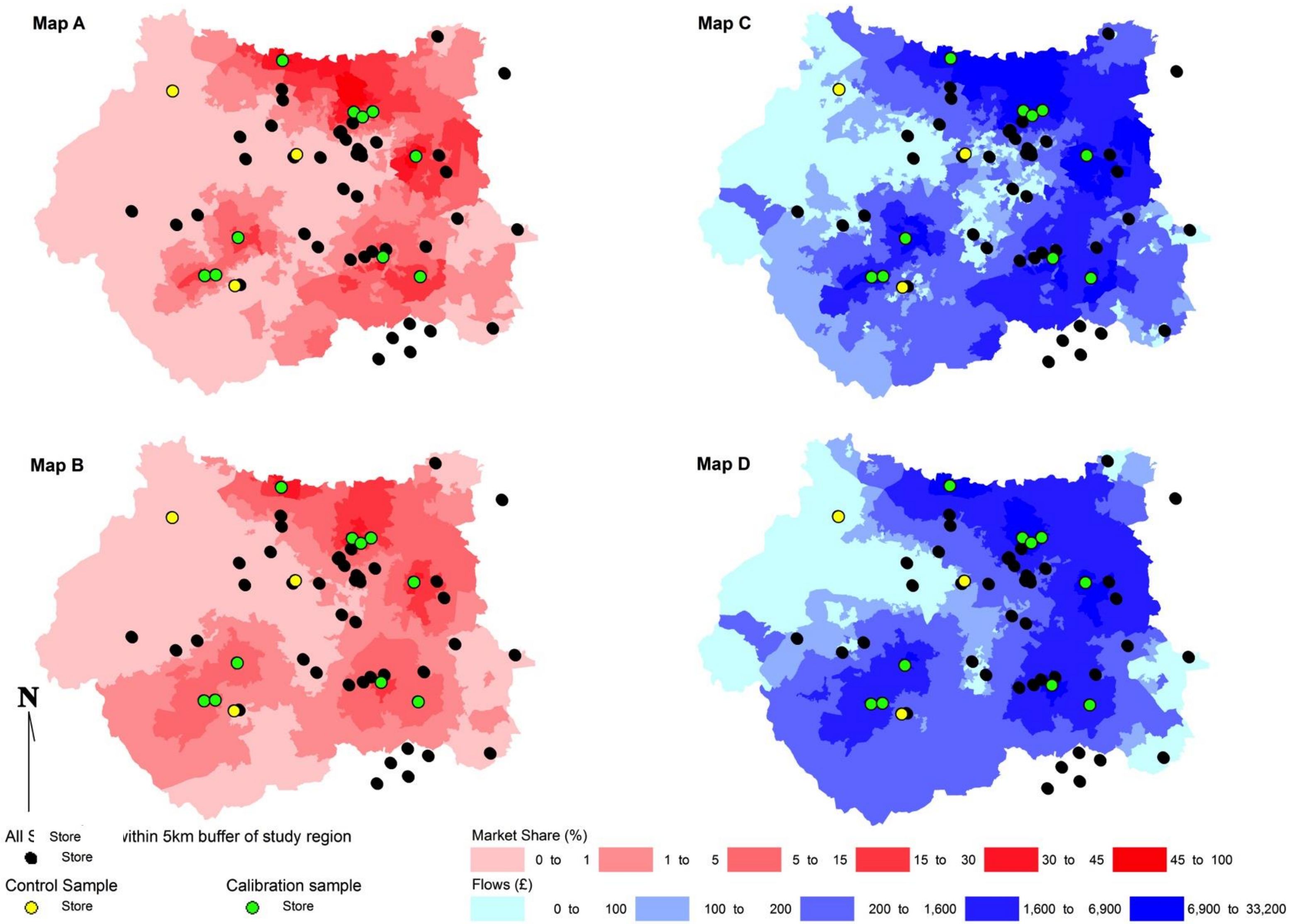


Table 6.8 – Showing predicted regional market share compared to observed values for retailers and their Goodness-Of-Fit following calibration on alpha in the SIM.

Observed regional market share and predicted market share for retailers operating in West Yorkshire, UK			
Brand	Observed Market Share % (*oms*)	Predicted Market Share % (*pms*)	Performance $^{pms}/_{oms}$
Aldi	3.37	3.08	0.91
Alfred Jones	0.001	0.001	1.00
Asda	19.68	19.2	0.98
Booths	0.48	0.46	0.96
Co-op	6.57	6.33	0.96
Costcutter	0.2	0.19	0.95
Farmfoods	0.15	0.15	1.00
Iceland	1.24	1.31	1.06
Lidl	3.34	3.47	1.04
Londis	0.21	0.22	1.05
M&S*	3.55	3.58	1.01
Morrisons	22.53	22.69	1.01
Nisa	0.18	0.17	0.94
Ocado	0.28	0.26	0.93
One Stop	0.14	0.13	0.93
Premier	0.41	0.4	0.98
Sainsbury's	15	15.09	1.01
Spar	0.13	0.12	0.92
Tesco	20.95	20.29	0.97
Waitrose	1.59	1.38	0.87
Goodness of fit for observed and predicated regional MS for all retail brands			
Correlation	0.999709946	R^2	0.999419977

* M&S was estimated at a regional level: this was derived on the assumption that the remaining portion of regional market share, which was not made up by all the other retailers i.e. the remaining Percent, was therefore more than likely M&S (M&S was estimated using the remaining market share as a proxy as it is not currently used as part of Kantar Worldpanel National market share data).

Table 6.9 – Alpha values for 2011 OACs and retailers calibrated via regional market shares data.

2011 OACs alpha parameter values for retail brands (weighted by 2006 OACs)								
Retailer	Rural residents	Cosmo-politans	Ethnicity central	Multicultural metropolitans	Urbanites	Sub-urbanites	Constrained city dwellers	Hard-pressed living
(Brand)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Aldi	1.0108	1.0070	1.0065	1.0079	1.0100	1.0100	1.0104	1.0098
Alfred Jones	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
Asda	1.0127	1.0098	1.0169	1.0175	1.0152	1.0153	1.0193	1.0188
Booths	1.0587	1.1348	1.0683	1.0560	1.0585	1.0553	1.0428	1.0445
Budgens	1.0000	0.9979	0.9924	0.9948	0.9987	0.9989	1.0004	0.9998
Co-Op	0.9400	0.9380	0.9324	0.9349	0.9387	0.9390	0.9405	0.9398
Costcutter	0.7400	0.7400	0.7400	0.7400	0.7400	0.7400	0.7400	0.7400
Farmfood	0.7700	0.7700	0.7700	0.7700	0.7700	0.7700	0.7700	0.7700
Iceland	0.9652	0.9637	0.9661	0.9654	0.9647	0.9647	0.9645	0.9648
Lidl	0.9617	0.9625	0.9678	0.9649	0.9613	0.9613	0.9607	0.9619
Londis	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300
M&S	0.9547	0.9831	0.9565	0.9533	0.9565	0.9544	0.9488	0.9494
Morrisons	1.0123	1.0081	1.0117	1.0125	1.0127	1.0127	1.0138	1.0134
Nisa	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300
Ocado	1.0387	1.1148	1.0483	1.0360	1.0385	1.0353	1.0228	1.0245
One Stop	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300	0.7300
Premier	0.7380	0.7380	0.7380	0.7380	0.7380	0.7380	0.7380	0.7380
Spar	0.7200	0.7200	0.7200	0.7200	0.7200	0.7200	0.7200	0.7200
Tesco	1.0292	1.0256	1.0253	1.0258	1.0269	1.0275	1.0252	1.0261
Waitrose	1.0087	1.0847	1.0182	1.0059	1.0085	1.0052	0.9927	0.9945
Sainsbury's	1.0156	1.0200	1.0109	1.0106	1.0138	1.0135	1.0066	1.0082

Table 6.10 – Current goodness-of-fit assessment of the SIM following alpha and beta calibration on flows and ATD

Current SIM GOF following alpha calibration and refinement of the beta parameter				
	ATD		Flows	
OAC	Beta	ATD^{O_g}	R^2	SRMSE
1 - Rural residents	0.315	1.00	0.76	1.94
2 - Cosmopolitans	0.58	1.01	0.53	4.98
3 - Ethnicity central	0.666	1.00	0.95	17.97
4 - Multicultural metropolitans	0.7	1.00	0.84	4.69
5 - Urbanites	0.5	1.00	0.88	2.37
6 -Suburbanites	0.47	1.00	0.86	2.7
7 - Constrained city dwellers	0.5	1.00	0.72	4.21
8 - Hard-pressed living	0.51	1.00	0.80	3.44
Overall GOF			0.76	3.4

6.3 Conclusions

In this chapter I presented a custom built disaggregated SIM using commercial data affording the opportunity to accurately account for observed consumer behaviour in the region. The updated methodologies indicated that following each stage of SIM development there was an increase in model accuracy, thus supporting the decisions for increased disaggregation of demand. The assessment of the model accuracy for revenue prediction through GOF statistics on flows shown previously (Table 6.10), reveals that while current modelled demand is behaving according to actual consumer behaviour, shown through the similar observed and predicted regional market shares and ATD (see Table 6.8 and Figure 6.4), the sample of stores that were used for calibration are still under-predicting (although some stores not used as part of the calibration sample are also over-predicting), demonstrated through the poor model fit for flows (in £). Monetary flows and store revenues largely remain lower than expected, based on observed consumer spending and observed store revenue. GOF assessment on flows following calibration of the alpha parameter at this stage gave a R^2 of 0.76 and SRMSE at 3.4, accordingly. While the overall evidence indicates that the model is more accurately replicating reality, in order to improve accuracy of predictions it is increasingly apparent that current demand estimates within store catchments are insufficient. Generally speaking, the supermarkets are well predicted, this is less so in the smaller convenience stores which are known to be supplied by localised foot borne trade, which are likely to be more sensitive to spatiotemporal variations. This is particularly noticed in the model for less densely populated areas such as the city centre, supporting the evidence of chapters four and five, noting that these areas often experienced a considerable spatiotemporal transition in demand over the course of a day. This evidence implies that consumers are likely spending on food and drink outside of the traditional grocery shop behaviour. Different demand types, following different behavioural and spatiotemporal patterns are likely alternative sources of revenue. This corresponds with demand being redistributed both spatially and temporally as shown in the earlier analyses. The result being that current estimates and model outputs offer limited insight. Therefore, it is likely that obtaining a higher level of accuracy, i.e. more realistic revenue predictions, can be achieved by incorporating spatiotemporal demand spend on food and drink in the grocery industry, as shown in early chapters, into the model. Spatiotemporal demand will initially be incorporated through the addition of workplace demand and is discussed in Chapter seven. It is reasonable to assume that this will improve demand estimates and improve model predictions, particularly in less residential and more urbanised locations, which were shown to fluctuate considerably. This will increase the accuracy of the model distributing expenditure throughout the model to improve the representation of daytime expenditure i.e. accounting for the differences of in and out flows of consumers during the daytime. The modelling and evidence provided in this and previous chapters indicate this need for temporal extension, which will likely improve model performance, reduce current areas of weakness and afford novel insight on grocery consumer behaviour through extended disaggregation of the population.

Chapter 7 – Development of a robust workplace demand layer

7.1 Introduction

The process of adding in work-based consumers is in itself not a novel idea within location analysis. However, the novelty of this project is in building robust workplace demand layers to account for the apparent large impact on temporal store sales that workplace demand represents by using the newly developed WPZ data. Thus this chapter presents the detailed methodology and impact upon SIM revenue estimations from incorporating temporal dimensions using a workplace population. It provides insight on the implications of spatiotemporal demand for location planners. The application of adding workplace data also represents a novel opportunity to undertake modelling of the workplace grocery market at a previously unprecedented level of detail, due to the newly developed and specialised datasets. In the previous chapter I demonstrated that a custom built SIM, in this case for West Yorkshire, using observed consumer behaviour from loyalty card data, results in a marked improvement in predictive accuracy over those derived solely from using residential census data. The model was shown to be more capable of replicating observed patterns of trade and consumer behaviour within the area. However, analysis of these extensions highlighted that while the model was capable of replicating behaviour from a residential population perspective, there was still a considerable level of inaccuracy more than likely caused by a lack of non-residential demand especially for convenience stores. It is known that residential populations are a poor representation of actual daytime populations due to the levels of variation in population movements, (both spatially and temporally) that individuals make at certain times of the day (Bell, 2015, Martin et al., 2015). This is particularly important when trying to model retail activity as shifts in populations result in a very different distribution of consumers and retail expenditure. Bell (2015) and Martin et al (2015) note that models which do not account for this, focusing on a residential population only, will therefore be limited in scope and accuracy. The redistribution of revenue caused by diurnal population movements can substantially affect store trade profiles as demonstrated previously in Chapter Five. This results in considerable spatial variations in revenue availability that modelling with a residential population is not capable of simulating alone. This potential for inaccuracy has been predominantly noted in the convenience market with stores located in low residential but high workplace population catchments (with significantly underperforming model-based estimates of revenue).

This chapter therefore aims to build on the previous model, linking the expenditure and revenue predictions of the residential consumer, (where we assume shopping trips originate from an individuals' home) with the redistributed expenditure of a daytime workplace population also included. Following the analysis of Chapters Four and Five, noting the temporal variations in trade, it is reasonable to assume that workplace populations represent a sizable source of revenue variation in the grocery market. Therefore, the initial spatiotemporal SIM extension incorporating temporally informed demand will be the addition of workplace (WP) consumers. This SIM, separately calibrated, will then be combined with the residential based model (see equation 7.4 below), demonstrating the processes for adding additional temporal dimensions within the SIM. Accounting for a more realistic daytime population allows for a far better expenditure distribution as well as recreating the spatiotemporal behaviour associated with daytime populations. This leads to an improved simulation of actual market conditions and improved revenue predictions.

It should be noted that in the past previous SIMs have been noted to account for WP uplift, although this has often been done in a far cruder manner in the past and often estimated from less robust data. The recent development of a purpose built and vastly more detailed WP dataset has been made available, therefore providing greater reliability to the demand estimates and a marked improvement in detail and accuracy. Subsequently, the aforementioned intended extensions offer not only novel insight, but also a potential level of detail and accuracy that has previously been relatively restricted in an academic setting. The framework of this chapter will be as follows. First the discussion will focus on the use of relatively new and purpose built WPZ data, briefly outlining on how the WP data were derived. This is followed by the development of the workplace based model, and model predictions will be estimated, accounting for temporal fluctuations caused by workplace demand.

7.2 Driving grocery sales: Workplace spending

As previously discussed in Section 2.7, consumer behaviour and spending are not only different throughout the day, but a considerable amount of daytime spending varies spatially, such as at an individual’s places of work. To clarify this argument Figure 7.1 identifies customers shopping at a city centre convenience store compared to their residential origins. It is clearly evident that many of the consumers shopping at the store do not use it as the ‘close to home’ type store. A possible explanation for this non-residential demand is commuters making regular trips to Leeds for work, or tourists and leisure shoppers who use the store due to its convenient locations. Likewise, if we consider the temporal trade profile (Figure 7.2) of the same store, also noting clear uplifts of trade throughout the day, such as at lunchtime or at the end of the ‘regular working day’. It is reasonable to assume these are linked to workplace shopping. Therefore, the existence of workday peaks in trade and ‘non-residential’ demand that can originate from outside a regular catchment area for a grocery store of that type and size, further supports the need to account for additional non-residential demand types and the decision to include a daytime population through workplace demand.

Figure 7.1 – Count of transactions at a city centre convenience store (red dot) originating at consumers' usual LSOA of residence.

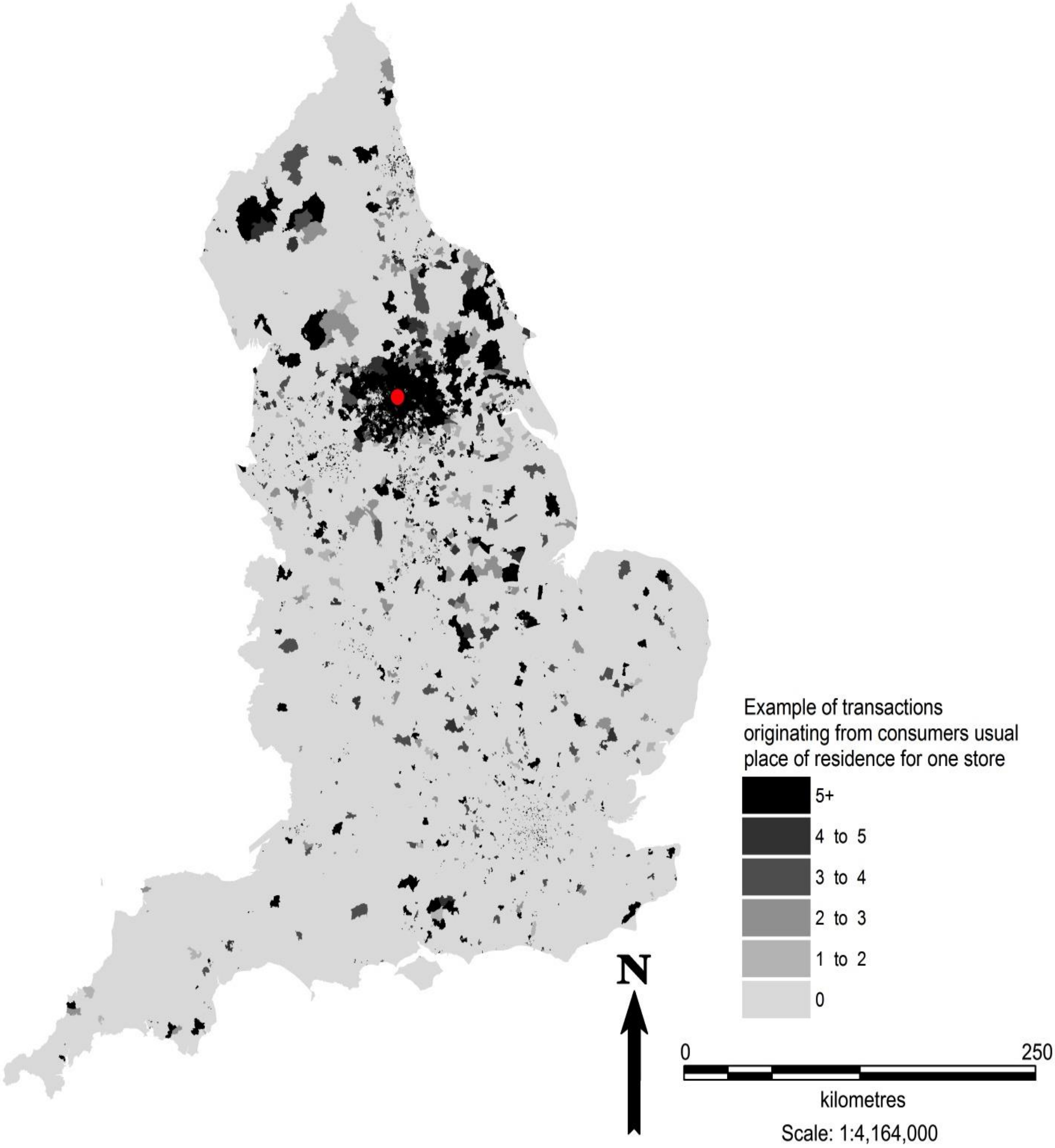
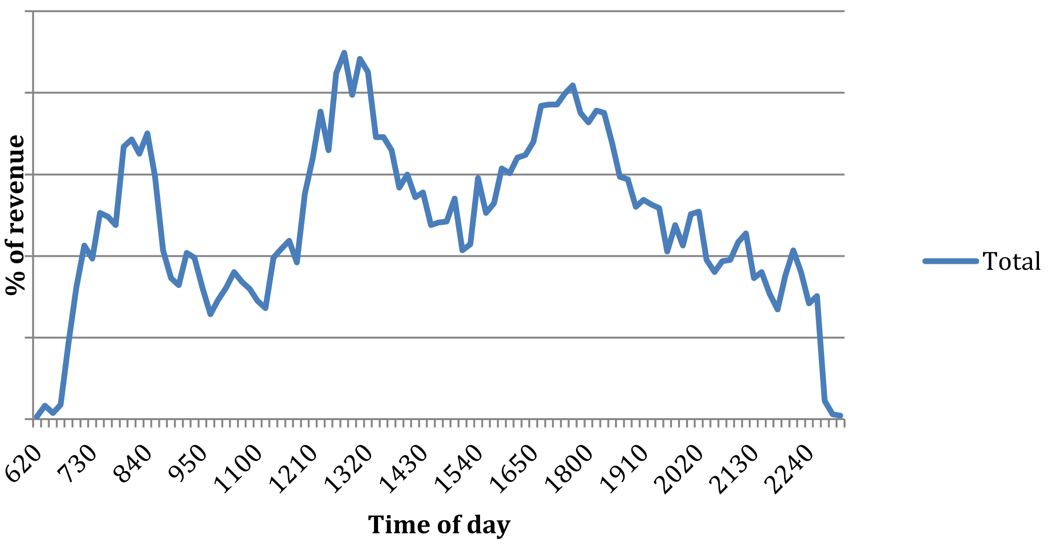


Figure 7.2 - Percentage of weekly store revenue at a city centre convenience store by time of day.



Although non-residential demand, such as workplace demand, remains somewhat under-researched (Berry et al., 2016), it is known to have a considerable impact upon the retail market and thus represents an important aspect of temporal behaviour and spending which needs to be considered. Evidence suggests that workplaces present a different population distribution to that seen in the traditional census data and in today's society the differences between these populations are typically large (Bell, 2015, Martin et al., 2013). Arguably, information informing the model must therefore come from an alternative source if a higher level of accuracy is to be achieved. The behaviour of WP consumers has been discussed earlier in Chapter two. However, in regard to WP shopping specifically within the grocery sector, further details are presented below.

Anecdotal evidence gained through industry representatives and location-based analysts, suggests that daytime spending equates on average to the value of £5 per person per week. This low figure is justified through industry market research and the assumption that not all individuals will buy lunch during the day and that others may in fact spend more. Secondly, the £5 value is applied weekly rather than as a daily expenditure because it is also assumed that those who do buy lunch, particularly in city centres, may choose from several lunchtime alternatives. For example, they may choose an alternative provider to a supermarket retailer. Industry representatives argue that while a total weekly 'daytime' spend on food was more likely to be around £25 per week, they estimate that 80% was spent at other service suppliers, such as fast-food retailers, coffee shops and takeaways. Therefore, if we assume that all food spending is accurately represented within the Living Cost and Food Survey (ONS, 2015a) and apply a £5 per person per week value to the working population in West Yorkshire, UK, the total workplace demand based on this rate is approximately £5.1 million, 8.4% of total weekly grocery expenditure. This expenditure has traditionally been allocated to residential addresses and not places of work, which are spatially very different. The shift in demand, discussed in greater detail in Section 7.3, demonstrates the considerable change in population and expenditure distribution throughout the day, which will arguably have a considerable impact upon store revenue patterns throughout the day. Additionally, WP consumers may adopt different shopping behaviour to a residential 'typical weekly' shop consumer. Proximity is argued as having the greatest influence in store choice for WP consumers, (more than other typical factors such as store size) with convenience of location and minimising the time taken, being major driving factors in destination choice (Schwanen, 2004). The development of the convenience store, in urban centres to cater for this consumer demand tallies with this assumption as one notes the considerable development and growth of this market in recent years (Berry et al., 2016, Birkin et al., 2017, Dashwood, 2013b, Hood et al., 2015, Kenhove and De Wulf, 2000). Furthermore, additional research conducted alongside industry professionals suggests that WP consumers are unlikely to travel further than 5 minutes (approximately 500m) to make food purchases within a city centre (Berry et al., 2016). This evidence, therefore, not only offers interesting insight into WP behaviour, but also is an important consideration for location modelling calibration. Store catchments will therefore be reduced, resulting in the need for higher beta values (controlling the deterrence of distance) as trade is driven predominantly by localised on foot sales.

To summarise, evidence suggests that traditional residential census data are insufficient for accurately predicting sales revenue in a SIM, particularly in relation to work-based consumers. In some instances, residential demand, typically considered representative of a night-time population, may be very low, whereas, in reality areas may be very densely populated throughout the day due to the presence of a workforce and are likely to be underestimated using residential only data (see section 7.3). Therefore, for SIMs to predict store revenue and consumer behaviour more accurately it is important to include a workplace orientated demand layer, otherwise groups of stores may be continually under-predicted due to the absence of the observed higher trading level during the day. The limitations and weaknesses of the residential based census for predicting workplace consumer behaviour, as in this instance, are widely reported (Berry et al., 2016, Birkin et al., 2010, Martin et al., 2013) and consequently the nationally recognised data for workplace zones (WPZs) were developed. Whilst previous location based modelling research notes that WP spending will improve model performance, it remains relatively under-researched, with even fewer studies incorporating the purpose built WPZ statistics recently released by the Office of National Statistics. Until models fully incorporate an improved daytime population data, they are likely to suffer from weaknesses and inaccuracy in these situations. The next section introduces the purpose built WP data. They will be used for development and customisation of the SIM in an attempt to refine the time-of-day fit, improving model predictions and the representation of temporal fluctuations within observed consumer demand.

7.3 Workplace zones

While traditional census based geographies such as OAs and LSOAs continue to provide detailed information regarding residential populations, as noted, they reflect a 'night time' population. A major limitation of these geographies and statistics is that in many instances highly residential OAs contain relatively small workplace populations, whereas OAs with low residential populations can often have high proportions of worker populations. Figure 7.3(a), below, illustrates this argument using a density map of Leeds City centre as an example. In the map the percentage of workplace population (in red) is much higher in the city centre compared to the residential population (in yellow). This further illustrates that traditional census data lacks relevance when accurately trying to represent temporal components of consumer demand, behaviour and spending in location based retail analysis. For example, daily commuting leads to a very different distribution of the population throughout the 'typical' working day and hence, model predictions will consistently fall short of observed revenue if models ignore transient populations. Figure 7.3(b) shown below, presents the distribution of the total WP population throughout the study region of West Yorkshire. The areas with the highest populations are (as expected) found in and surrounding major towns and cities in the region such as Leeds, Halifax and Bradford. A count of 116 WPZs has a total population of over 1000 workers. The top three WPZs (all above 6000 workers) are located in Leeds (x2) and Halifax; the former classified as 'administrative centres' and a 'major hospital' and the latter 'a regional business centre'. Contrastingly, the corresponding residential populations for these major WP centres are much lower. The maximum residential population of the three (the Halifax location) is 1157 residents, demonstrating the potential levels of temporal fluctuations in demand caused by population movements. Outside of the region's major urban sites there are several areas that stand out on Figure 7.3(b); one in the north of the map (the site of the Leeds-Bradford Airport), one in the South of the map (explained by an industrial estate with large warehouses for national companies) and in the North-East of the map (which is a combination of HMP Wealstun, The British Library and a large business estate). Table 7.1 shows that Leeds is the largest employment zone in the area with approximately 40% of West Yorkshires workers located in that area, and manufacturing and distribution are the largest employment sector in the region employing 23% of all workers.

Table 7.1 - Breakdown of workplace populations (%) by West Yorkshire administrative zones and Classification of Workplace Zones (COWZ)

	Bradford	Calderdale	Kirklees	Leeds	Wakefield	Totals
Manufacturing and distribution	4.3	2.0	4.2	7.3	5.5	23.3
Metro suburbs	4.3	0.2	1.7	4.5	0.2	11.0
Retail	2.1	1.3	2.2	4.5	1.7	11.9
Rural	1.3	1.0	1.8	1.7	1.7	7.4
Servants of society	3.4	1.3	2.7	6.8	2.3	16.5
Suburban services	3.2	2.2	3.4	6.1	2.8	17.7
Top jobs	2.1	1.1	0.1	8.7	0.2	12.3
Totals	20.6	9.1	16.1	39.6	14.6	100

Figure 7.3(a) – Dot density plot of household (left) and workplace (right) populations (% of) for Leeds city centre showing spatiotemporal transition between night and day time population.

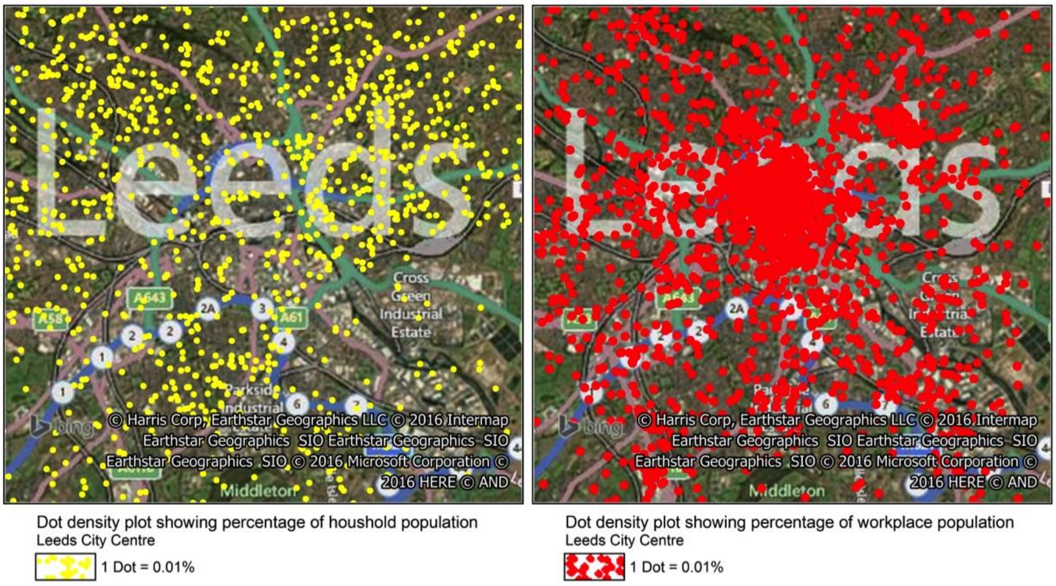
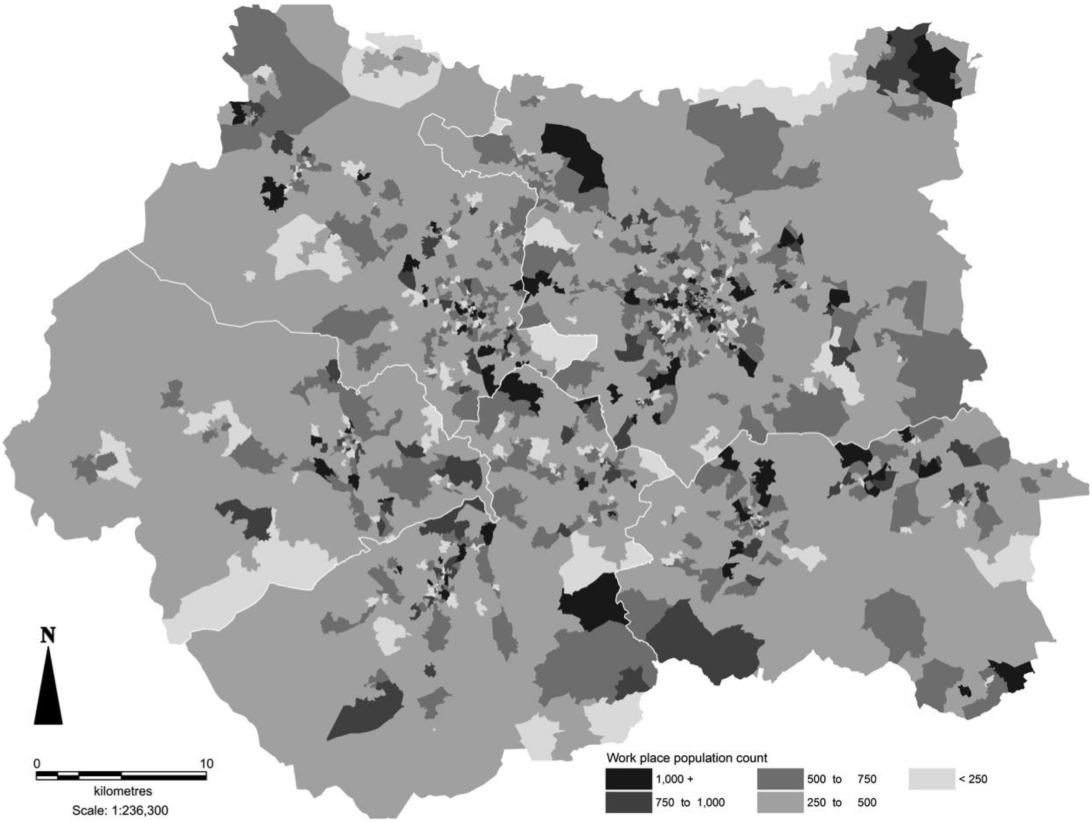


Figure 7.3(b) – A map showing the total populations of workplace zone throughout the West Yorkshire region.



Consequently, following extensive public consultation and academic involvement, the ONS released a new workplace based geography in 2013, termed 'workplace zones' (WPZs) (Berry et al., 2016, Martin et al., 2013, Mitchell, 2014) in order to overcome these issues. WPZs were designed to supplement residential geographies and as such were fashioned from OAs and nested within MSOAs. Each WPZ is contained within a single MSOA (Mitchell, 2014, Martin et al., 2013). Data related to counts of workers in each OA were generated through census records and applied with a series of constraints and thresholds used to inform their design and manage anonymity. Following the outlined framework in the ONS report, census data were used to generate individual WPZs through a process involving merging and splitting existing OAs: the full details of how they were created can be seen in the ONS report (Mitchell, 2014). In short, a minimum threshold of 200 workers and at least three workplace postcodes was set as a strict constraint to conform to confidentiality requirements. In some instances existing OAs had very small working populations and so these were merged to form much larger WPZs, such as in very rural locations. On the other hand, where possible, a maximum of 625 workers was set (although this was less strict and no maximum restriction on workplace postcodes was set). On the upper thresholds, the choice to only stipulate a maximum threshold was due to many OAs exhibiting workplace populations well over 625, as well as containing many more workplace postcodes. This was due to the fact that several PO boxes (pertaining to large workforces) were concentrated at a single point (Berry et al., 2016, Mitchell, 2014). Where an OA exceeded this maximum threshold, OAs were subdivided except in cases where this would lead to a single employer representing the entire workforce of a WPZ. To avoid disclosure of any one employer, or a single workplace postcode, WP population counts were capped at 100 workers per employer. This was to ensure individual employers were not discernible in the data. As such, in low residential and high industry locations, (for example the City of London) even though a single OA was subdivided into 115 WPZs, the mean WP count is still over 1000 (Berry et al., 2016, Mitchell, 2014). A total of 53,578 WPZs were created from the 181,408 OAs. Many large OAs (which contained low residential populations but a high WP population) were split and subdivided, and similarly many small OAs, with high residential populations but low WP populations, were merged together to abide by the set constraints. The final outcome was an average of one or more OAs per WPZ observed (Mitchell, 2014). Within West Yorkshire this represented an increase in the total number of zone boundaries, from 1388 LSOA boundaries to a total of 2096 WPZ boundaries.

7.4 Building a customised Workplace SIM

Similar to the residential based SIM in section 6.2.3, a customised and disaggregated SIM specifically designed to operate using workplace demand and geographies was developed. The equation is written below:

$$S_{kj}^{bf} = A_k O_k^z W_j^{\alpha bf} \exp^{-\beta c_{kj}} \tag{7.1}$$

where:

S_{kj}^{bf} represents predicted expenditure between workplace zone k and store j by brand b and store format f. Store format f differentiates between store formats of major retailers based on whether store j is a supermarket or a convenience store.

O_k^z represents total grocery expenditure in workplace zone k

where:

$$O_k^z = T_k^p \gamma \tag{7.2}$$

and;

T_k^p is the total workforce population in workplace zone k (including those working from home or near their home, those travelling to work from another location within the model and those traveling to work from outside of the model).

γ represents the value of expenditure per person spent in the grocery sector at the workplace

$W_j^{\alpha bf}$ is the measure of attractiveness of store j raised to the power of α^{bf} $where$ α^{bf} is a power function determining the importance of attractiveness variables for store j by brand b and by store formant f.

$exp^{(-\beta c_{kj})}$ is a distance deterrence factor impacting the distance travelled between origin k and retail destination j.

A_k represents the balancing factor controlling competition in the model ensuring that all demand from origin k is distributed to stores within the study area and is expressed below in the following equation

$$A_k = \sum_b \frac{1}{\sum_j W_j^{\alpha^{bf}} exp^{(-\beta c_{kj})}} \tag{7.3}$$

The decision to undertake store format disaggregation was based on an evaluation of the shortcomings highlighted in the outputs of the residential SIM. In particular the SIM fell short in revenue predictions and trade intensity across nearly all convenience stores. Figure 7.4 (below) illustrates the accuracy for revenue predictions and trade intensities (£/sqft) for supermarkets and convenience stores prior to the redistribution of workplace expenditure. Although supermarkets were also under predicting overall (figure 7.4 (a) and (b)), predictions were largely far more accurate with an average accuracy level of 90%. The level of accuracy of convenience stores (figure 7.4 (c) and (d)) was considerably lower with an average revenue prediction rate of only 55%. Although this problem is partially a demand issue for stores located in WPZs, which this chapter aims to address, many convenience stores in residential suburbs were still under-predicting. This evidence supports the assumption that convenience stores required a separate brand attractiveness parameter in order to optimise their predicted trade intensity within an acceptable level of observed conditions, meaning they needed to be far more attractive than they currently were in order to trade properly. The current measure of a store's attractiveness is floorspace size (sq./ft.). This therefore has more of an impact on the revenue predictions of convenience stores in the model because they are seen as less attractive due to their smaller size. In reality, stores often trade at much higher intensities when compared to their larger counterpart stores, particularly, in city centre locations because of their considerably increased ease of convenience and the levels of footfall at certain times and therefore require a higher alpha value. It should be noted that the processes of disaggregating brand format were undertaken on both the residential and workplace model to ensure synergy between model predictions and supply-side behaviour.

Figure 7.4 - (a) Observed and predicted revenue for supermarkets

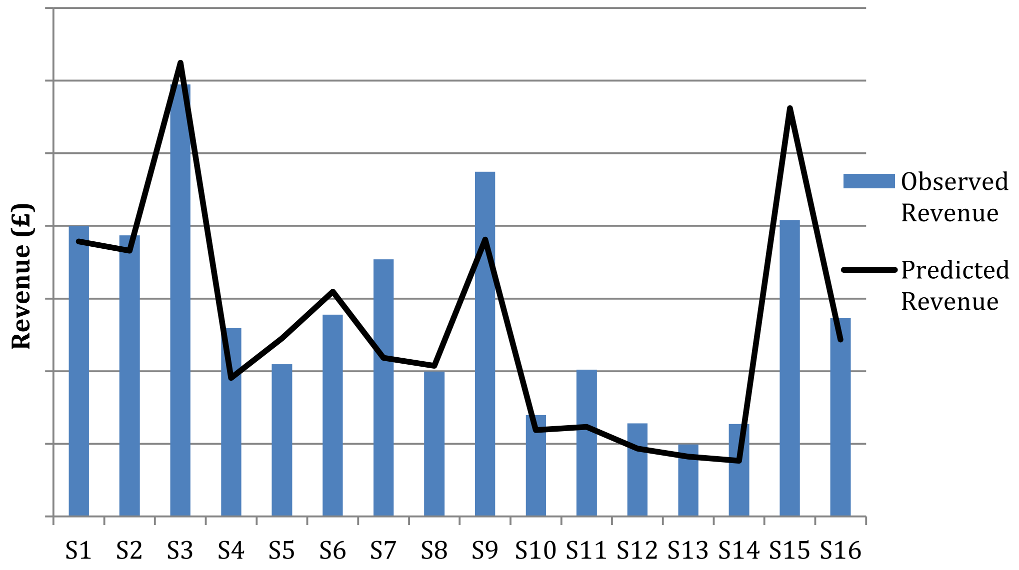


Figure 7.4 - (b) Observed and predicted trade intensity for supermarkets

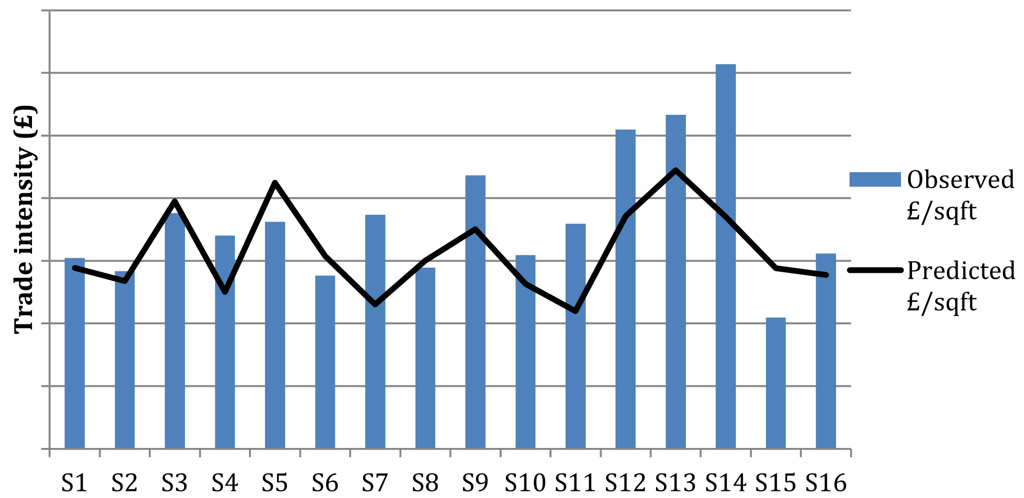


Figure 7.4 - (c) Observed and predicted revenue for C-Stores

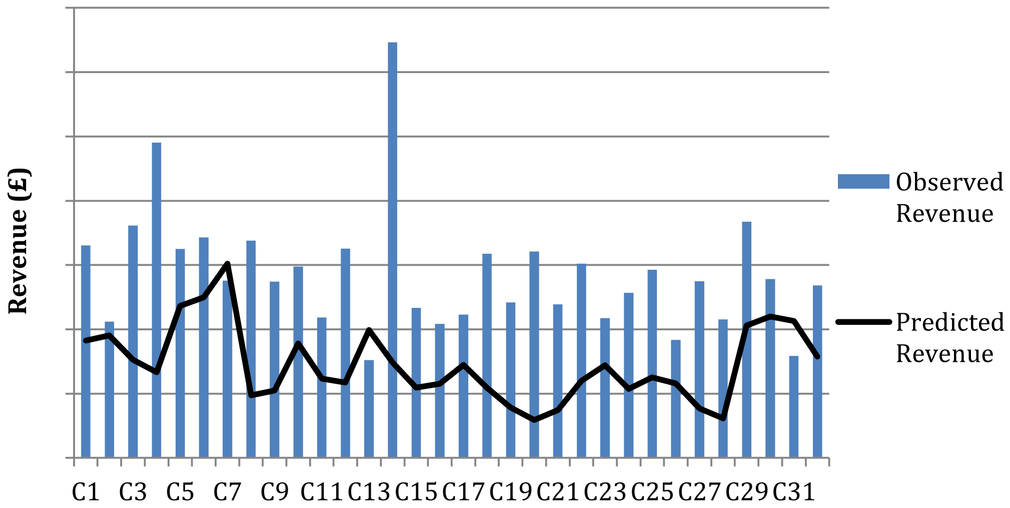
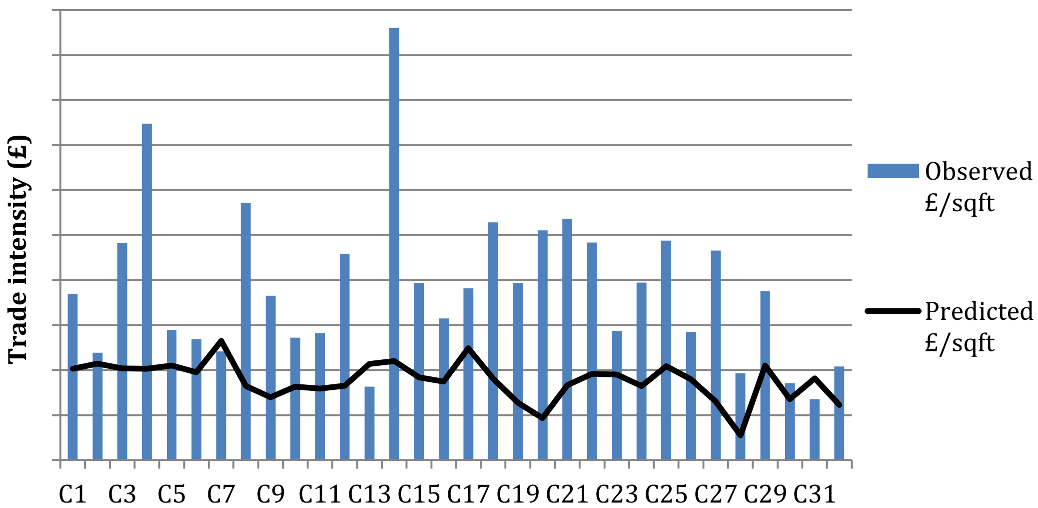


Figure 7.4 - (d) Observed and predicted trade intensity for C-stores



Demand estimates for per person workplace spend (γ), as noted in section 7.2 were informed through industry insight and was set to £5 per person per week. It should be noted that few previous studies, building such a comprehensive workplace zone SIMs, have been informed using observed consumer trading and there is little supporting material to suggest an alternative value. However, to ensure that £5 figure was the most 'optimal' WP demand value, and to avoid artificially adding additional expenditure or over-fitting the model, a series of different workplace expenditure values and scenarios were tested as part of the calibration process. First, a range of expenditure values were tested. Whilst the higher expenditure figures boosted the accuracy of stores predicting very badly, the stores with previously more accurate predictions were skewed and over-performed significantly. Expenditure values were then tested in a series of scenarios: (i) WP expenditure was already accounted for in The Family Spending Survey and was therefore redistributed from residential addresses to workplace addresses; (ii) WP expenditure was entirely new and represented additional grocery demand and was allocated to the relevant WPZ; (iii) WP expenditure was a mix of new and existing expenditure and was allocated to relevant WPZs as a mix of new and redistributed revenue (the weighting of new and existing revenue was also tested). Following the what-if scenario tests, the most realistic revenue outputs corroborated with the industry derived value of £5 per person for the workplace population.

Calibration of the brand format disaggregated alpha parameter was firstly undertaken on the daytime residential population. It was deemed the most appropriate solution to calibrate these using the loyalty card data because little data pertaining to WP behaviour and attitudes towards grocery brands were available. By using the daytime residential SIM, it is possible to disaggregate and calibrate alpha according to brand and store format using an estimate of observed regional market shares, calculated in table 6.7, and an estimate of predicted regional market shares attained using a similar method. Following the optimisation of these market shares through adjustments made to attractiveness it was possible to disaggregate alpha by brand and format as well as household, while maintaining the optimal regional market share for each retailer in the model to ensure a realistic share of grocery demand and avoid dominance of one retailer or store type. These parameters were then reproduced within the WP SIM. Work based consumers, as discussed, tend to seek stores which are convenient to their current workplace location with factors such as ease of access and distance acting as a primary stimulus on decision making (as previously discussed in Chapter two). Beta was realigned to fit within the constraints implied by academic and retail industry opinion, both noting that catchments are typically much smaller for stores catering to WP consumers. Due to the lack of observed consumer data relating to the patterns of WP behaviour, beta was adjusted so that flows fell within a loose upper threshold of 1km. The constraint of beta was softened from industry evidence, increasing the 500m threshold noted by Berry et al (2016) due to retail markets in London typically having smaller catchments and more concentrated catchments when compared to the rest of the UK. The outputs detailing predictive performance following this refinement of the time-of-day fit for SIMs can be seen below in section 7.5. These results demonstrate a continued improvement in model performance, particularly in convenience store accuracy following the diurnal distribution of WP populations and resulting continued calibration.

The resulting model, combining revenue predictions from both the residential and WP SIM (representing an initial refined time-of-day fit) is expressed below through the following equation.

$$S_{ij}^{gbf} + S_{kj}^{bf} = (A_i^g O_i^{rg} W_j^{\alpha^{gbf}} \exp^{-\beta^g c_{ij}}) + (A_k O_k^z W_j^{\alpha bf} \exp^{-\beta C_{kj}})$$

(7.4)

where:

S_{ij}^{gbf} represents predicted expenditure between origin i and store j by household classification type g and store brand b and store format f. Store format f differentiates between store formats based on whether store j is a supermarket or a convenience store.

S_{kj}^{bf} represents predicted expenditure between workplace zone k and store j by brand b and store format f.

O_i^{rg} refers to total residential grocery expenditure remaining in origin zone i.

O_k^z represents the total workplace grocery expenditure in workplace zone k.

$exp^{(-\beta^g c_{ij})}$ a distance deterrence factor, incorporating changes in behaviour for household type g, impacting the distance travelled between origin i and retail destination j.

$W_j^{\alpha^{gbf}}$ is the measure of attractiveness of store j to the power of α^{gb} $where$ α^{gb} is a power function influencing the importance of the attractiveness parameter for store j by household classification type g and store brand b and store format f.

and;

$$O_i^{rg} = O_i^g - O_i^z$$

(7.5)

O_i^g equals total grocery expenditure in zone i by household g

O_i^z equals total grocery expenditure leaving zone i to be spent at the workplace

and;

A_k is expressed above in equation 7.3

A_i^g representing the balancing factor controlling competition in the model ensuring that all demand from origin i by household classification g is distributed to stores within the study area which is calculated through the extension shown in the following equation:

$$A_i^g = \sum_b \frac{1}{\sum_j W_j^{\alpha^{gb}} exp^{(-\beta c_{ij})}}$$

(7.6)

7.5 Assessment of modelling outputs following temporal extension

This section aims to demonstrate that through the initial incorporation of a robust temporal workplace demand layer, the models' time-of-day fit (by accounting for temporal components of demand) has increased, presenting a far more realistic account of daytime population distributions. Although only an initial temporal development, this demonstrates the benefits for predictive capabilities that adding in temporal demand fluctuations can provide to SIMs. Newing et al (2013a, 2014b), incorporating seasonal demand fluctuations, demonstrated that revenue predictions were improved when temporal dimensions were accounted for (see below). Similarly preliminary results produced far more store predictions within an accuracy threshold of 10% +/- of observed revenue, further demonstrating the importance of continued time-of-day fit modifications through the incorporation of temporal demand fluctuations.

Figures 7.5 and 7.6 below detail the current revenue predictions from the extended model. Revenue predictions are compared to observed revenue and likewise observed trade intensities for both supermarkets and convenience are compared, providing an assessment of the predictive capacity of the new model. The biggest improvements in revenue predictions are observed in the convenience format stores. This is due to the fact that many convenience stores are located in sites supplied with higher daytime/non-residential populations (Birkin et al., 2017) which were not accounted for in the previous model. Comparatively, Figures 7.4 (c) and (d) and Figures 7.5 (a) and (b) and 7.6 respectively, demonstrate these improvements, providing credibility to this argument. It is clear that the disaggregation of demand and improved time-of-day fit of population distributions has overall had a positive impact upon the accuracy of the revenue predictions (see table 7.1). The improvements of this approach demonstrate the success of combining separate models that are temporally informed, resulting in better modelling opportunities that new data and continued research can provide. Generally speaking, revenue predictions improved following this process. However, further temporal refinement will be necessary as some stores remain outside an acceptable accuracy threshold and some stores now significantly over-predict, suggesting the need for additional refinement.

Figure 7.5 - (a) observed and predicted revenue following incorporation of a workplace population for supermarkets

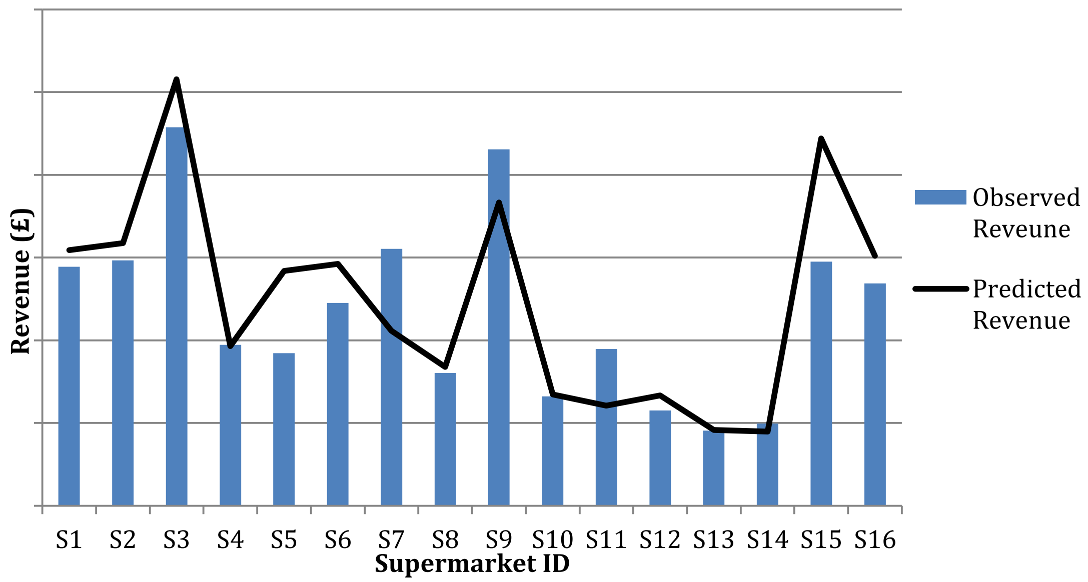


Figure 7.5 - (b) observed and predicted revenue following incorporation of a workplace population for convenience stores

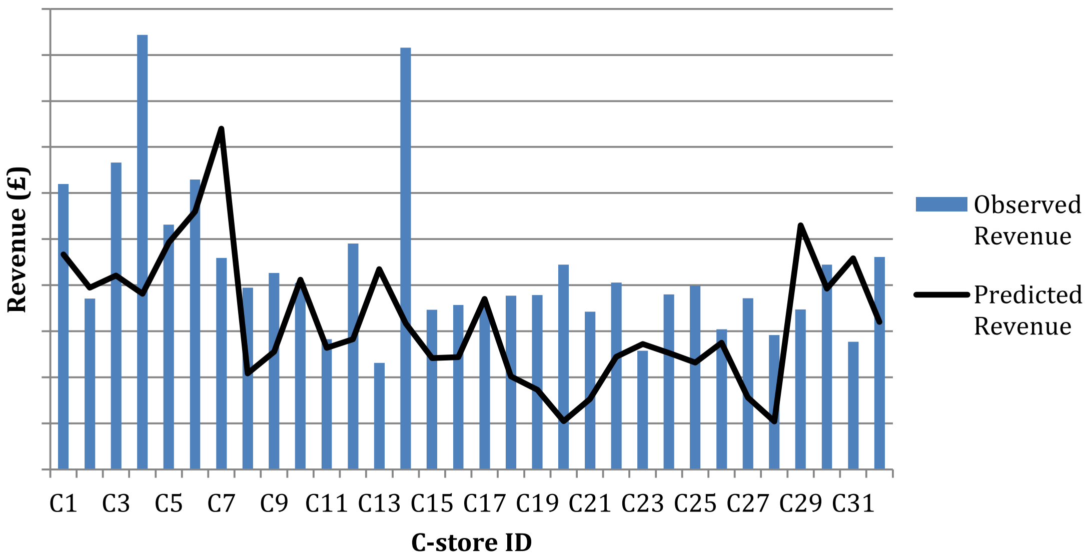
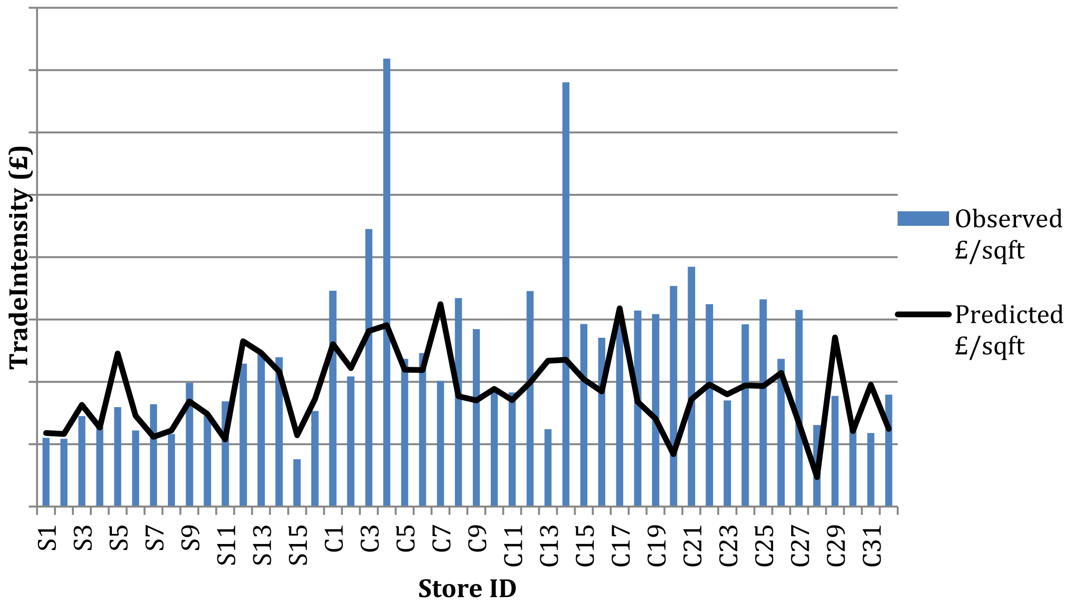


Figure 7.6 - observed and predicted trade intensity (£ per sqft) resulting from the current model build



While model predictions have generally improved following these temporal extensions, when looking at accuracy there are still a significant number of poorly predicting stores within the model. This is predominantly associated still with convenience stores. This provides an interesting insight into convenience trade. While revenue estimates in the convenience market were noted to be those most affected by the incorporation of a workplace demand, (demonstrating the greatest sensitivity to a workplace population as a source of revenue) these were typically still the worst performing stores in terms of predictive accuracy in the model. This evidence suggests that a far more complex interaction between supply and demand exists and that to fully predict store revenue, particularly in these locations, additional dimensions of demand might need to be included. In contrast, the incorporation of a workforce population had a smaller impact upon supermarkets. In any case supermarkets were far more in line with observed revenue levels, resulting in a higher rate of predictive accuracy. If we compare the accuracy of this extended SIM with the night time model (shown below in Table 7.2), it is clear that temporal demand is not only an important factor for revenue estimation in the grocery sector but that it is important for understanding how to improve the capacity of location based tools.

Table 7.2 - Count of accuracy rates (%) from current extended model and night time model (residentially based).

	Night time model			Refined Time-of-day Model		
% of Revenue	All stores	Supermarket	Convenience	All stores	Supermarket	Convenience
Count <50	20	1	19	7	0	7
Count 50-60	7	4	3	4	0	4
Count 60-70	7	1	6	7	2	5
Count 70-80	6	6	0	3	0	3
Count 80-90	4	2	2	2	1	1
Count 90-100	2	2	0	6	2	4
Count 100-110	0	0	0	9	5	4
Count 110-120	2	0	2	4	4	0
Count 120-130	0	0	0	0	0	0
Count 130-140	0	0	0	0	0	0
Count 140-150	0	0	0	0	0	0
Count > 150	0	0	0	6	2	4
Average	57.07	70.86	50.17	88.95	106.02	80.41

A comparison of accuracy for ‘all stores’ between the night time model and refined model indicates that there has been a significant improvement in accuracy levels. Average ‘accuracy’ has risen by over 30%. Table 7.2 highlights that there are fewer stores under-predicting following the development of a robust workplace demand with a total of 15 stores predicting within a 10%+- threshold, compared with the night time model, which had a total of 96% of stores predicting outside this 10% threshold. It should be noted that while the refined model has improved predictive accuracy overall, in both models convenience stores performed the least well, representing the majority of stores that were under-predicting. The improvements on accuracy levels between the two models have been illustrated below (Figure 7.7 and 7.8). The histograms and the trend lines demonstrate the shifts in accuracy experienced between supermarkets and convenience stores when incorporating a temporally informed demand layer. As noted, in terms of proportional revenue, the impact of a workforce population has had less of an impact on total supermarket revenue predictions: although total sale increases were proportionally small, many of the stores are now predicting better, shown via the shift of the trend line. Likewise, convenience stores demonstrate a general uplift in predictive accuracy. However, as previously noted convenience stores revenues were observed to be far more sensitive to the impact of temporal demand, demonstrated by the sizeable increases in the proportion of total revenues (see Figure 7.4c and Figure 7.5b). The actual percentage increase to total revenue following the incorporation of a workforce population for supermarkets was 8% and over double that for convenience stores at 16%. The proportion of stores predicting within a 10% threshold were still however far less than supermarkets following the incorporation of a workforce population. It is reasonable to assume then that additional temporal components of demand clearly have impacts on both supermarket and convenience store revenue in regards to location modelling. However, it is likely that additional temporal components and accurately managing temporal dimensions in the grocery market will have a far greater impact on convenience store predictions within SIMs.

Figure 7.7 - Counts of revenue prediction accuracy for supermarket stores.

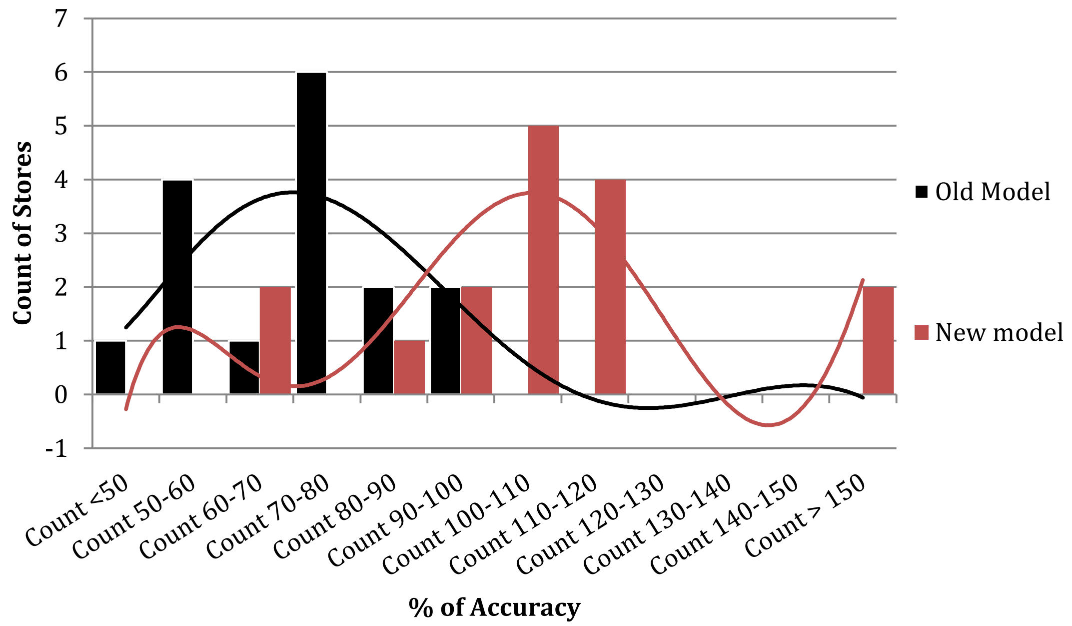


Figure 7.8 - Counts of revenue prediction accuracy for Convenience stores.

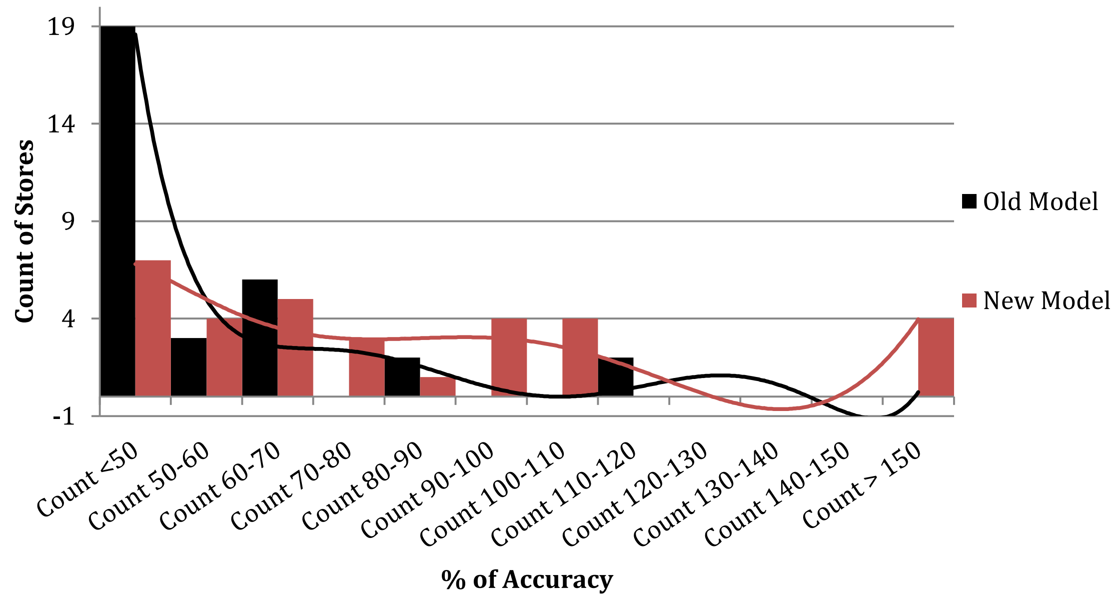
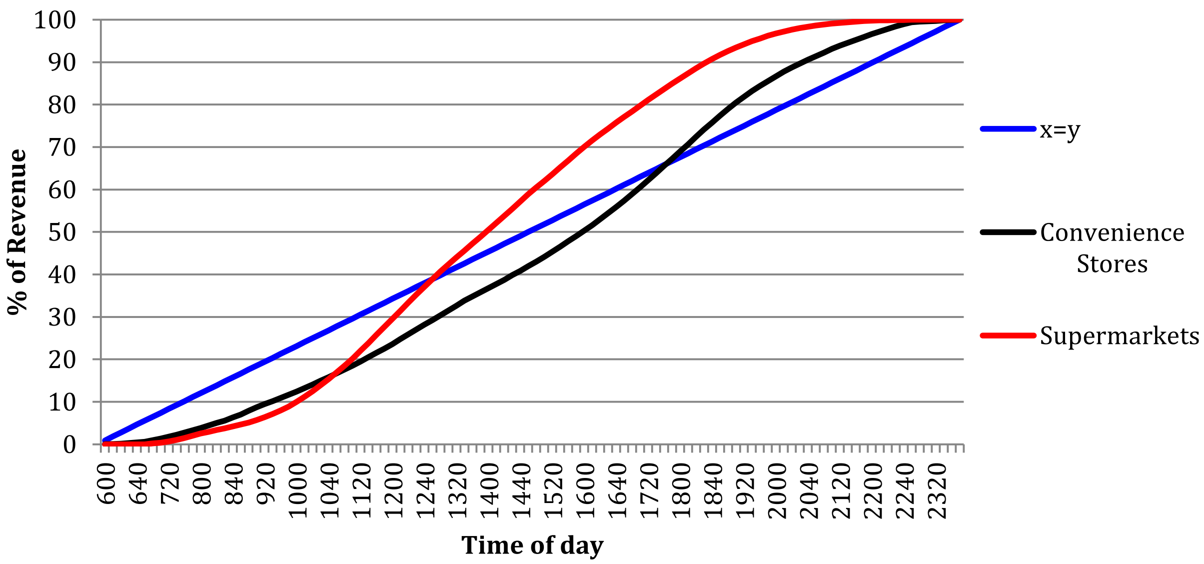


Figure 7.9 shown below reveals that 80% of supermarket revenue is made over a six hour period, whereas on the other hand 80% of convenience store revenue is made over 9.75 hours. Convenience store revenue is comparatively more varied over time than supermarkets (as

demonstrated in chapter five), which is likely due to demand in these locations being more temporally variable throughout the day compared to the more relatively stable demand at supermarkets. The evidence offers further justification for the need to accurately account for temporal fluctuations. The findings presented in Figure 7.9 are consistent with other major grocery retailers operating in the UK. A report on Waitrose convenience stores published in the Guardian newspaper describes how 40% of their revenue is generated after 17:00 (Smithers, 2015), a figure that is well in line with the findings of the analysis of the commercial partner's convenience store revenues. It is reasonable to suggest that in achieving a high level of accuracy for the partner's convenience stores (through calibration of both supply and demand temporal components), this will have similar results for other retailer brands in the SIM, ensuring a realistic representation of the grocery market. Therefore, the development of a robust temporal demand layer will not only improve overall model performance, but it will represent a marked improvement in the commercial sectors' ability to predict store revenue in the convenience market accurately, which anecdotal evidence notes is typically difficult to achieve (Hood et al., 2016).

Figure 7.9 - The cumulative percentage of revenue over time for Supermarkets and for Convenience Stores.



As noted in Table 7.2, although this initial temporal extension of the SIM helped to improve model accuracy overall, numerous stores are still poorly performing which can logically be taken as an indicator for the model's performance across all retailers. Only three of the stores under-predicting were supermarkets. It was also noted that stores linked to major transport sites, e.g. train stations, also under-performed by a significant margin. Of these stores, four were located on the boundary of the model and will likely continue to under-perform irrespective of

SIM extension. The stores currently over-predicting tended to be located in larger town centre sites such as *major town centre* and *major city centre* location type classifications and were more often than not supermarkets with an average sales area of 25,000 square feet. It is important to therefore look for further model refinements in the next chapter.

7.6 Conclusions

This example of temporal extension accounting for fluctuating demand, in this case workplace populations, appeared to provide a better representation of the intricacies of expenditure distribution and grocery store performance, demonstrating that spatiotemporally informed extensions to a SIM improving the time-of-day fit resulted in better revenue predictions. While the resulting model is not fully accurate compared to known data, the overall improvements further highlight the need to undertake this research. This view has likewise been expressed within previous academic research relating to SIMs, location based modelling and consumer behaviour, noting that accounting for supply and demand side-temporal factors, which are noted to have considerable impact on store revenue, will potentially result in far more accurate SIMs (Hernandez, 2007, Newing et al., 2014b). The outputs presented in this chapter have shown that generally the initial step of adding in workplace demand has resulted in a more accurate account of diurnal population distribution resulting in a general improvement in predictions. However, the existence of stores that are significantly under and over-predicting emphasises that the relationship between demand and supply is far more complicated. Accurately accounting for temporal distribution of customers, as well as modelling new demand types, will be an important step for achieving a substantial advancement in location based modelling predictive tools. As noted, in chapters four and five stores in major urban and city centres were observed to experience high levels of demand fluctuation during the day and exhibited corresponding temporal revenue profiles, which appear to be linked to the impact of workplace demand. In these instances, the predicted revenue has increased, helping to improve the overall accuracy of the model but also demonstrating the impact of improved demand modelling, via spatiotemporal disaggregation, at a localised store level.

In the next chapter, further temporal refinements to the SIM, such as school and university demand layers (drawing on the findings previously discussed in this thesis) are developed, as well as discussing the corresponding behavioural elements of consumers. The incorporation of drive time data between origins and destinations presents a more realistic representation of actual origin-destination interactions, replacing straight-line distance. This will be followed by a review of model performance, in addition to a detailed discussion of the strengths and weaknesses of a spatiotemporally informed model. Analyses of the areas that the model works, and does not work, and understanding why, represents an important stage in model refinement, as well as providing insight into different demand types and our understanding of their behaviour.

Chapter 8 - Daytime Spatial interaction model: Accounting for spatiotemporal demand fluctuations

8.1 Introduction

The previous chapter demonstrated that through the incorporation of temporally informed, and thus a more representative demand model, store predictions were generally improved. This chapter aims to further extend the number of the temporal components and demonstrate the importance of spatiotemporal fluctuations to retailers further through refinements in temporal and demand type disaggregation for modelling. In order to more accurately represent actual consumer travel behaviour, drive time data was incorporated in the model, replacing Euclidean distance used in the previous model iterations. The partner organisation provided the drive time data. This details the shortest drive time (in minutes) and corresponding drive distance (in kms) along road networks from all UK LSOAs to each LSOA in West Yorkshire. The beta values in the model were then recalibrated for each household OAC to ensure a representative distance deterrent was applied. The goodness-of-fit (GOF) for predicted and observed flows, based on loyalty card transactions in the residential model, is shown below in Table 8.1. The current GOF, following these drive time and workplace demand extensions demonstrates a slight improvement in the model's predictive accuracy, compared to Table 6.10. A limitation of this data is that store locations as a destination were defined by the LSOA centroid they were entirely within. Thus, a drive time from the same LSOA was initially cited as zero minutes. This poorly represents the actual drive times and could overly bias stores as there is no cost incurred to reach that particular destination. To reduce this problem, a minimum drive time of two minutes was added to LSOAs with a zero drive time value.

Table 8.1 – Goodness-of-fit assessment of the SIM following incorporation of drive time (in minutes) data.

	Flows	
OAC	R^2	SRMSE
1 - Rural residents	0.71	2.01
2 - Cosmopolitans	0.96	1.37
3 - Ethnicity central	0.91	23.19
4 - Multicultural metropolitans	0.79	5.13
5 - Urbanites	0.91	1.73
6 -Suburbanites	0.83	2.30
7 - Constrained city dwellers	0.72	4.09
8 - Hard-pressed living	0.74	4.49
Overall GOF	0.80	3.26

8.1.2 Store attractiveness measure

Before incorporating an increased suite of temporal components, it was important to ensure that the current model, which is representative of standard industry practice regarding demand disaggregation, was as realistic as possible. While traditionally the attractiveness of a store has been a measure of store size, the growing complexity of retail markets encourages the development of additional attractiveness terms within a SIM, which are often necessary for accurate store predictions (Birkin et al., 2010). The effect of spatiotemporal demand fluctuations on store revenue (chapters four and five) is a prime example of the complex market conditions experienced within grocery retail. As previously indicated in chapter seven, convenience stores tend to under-predict and supermarkets to over-predict in conventional models. Thus, a floorspace only attractiveness measure can logically appear potentially too simplistic and will bias flows to stores at each end of the floorspace scale, e.g. small c-stores and super large supermarkets. Birkin et al. (2010) suggest that floorspace may be only one attractiveness component. They suggested a series of potential attractiveness measures, namely ‘floorspace, store brand, store maturity, agglomeration and store/centre performance’ to account for increasing complexity in destination choice within retail models. In this research floorspace and store brand (by household type) were used for the attractiveness term. Store brand has been comprehensively calibrated using observed market shares and disaggregated by retail brand, store format and consumer type. This approach was adopted following similar previous successful applications (Newing et al., 2014b, Thompson et al., 2012). Further disaggregation was not possible as additional data was not available making this beyond the scope of this thesis.

8.2 Spatiotemporal demand

Temporal extensions to the model were implemented through the use of temporally informed demand layers, derived from the analyses in chapters four and five. Using empirical data relating to volume and the locality of each demand type, a series of novel demand layers were developed for West Yorkshire and incorporated into the SIM. The previous chapter demonstrated the impact of temporally informed demand layers, using workplace demand for initial development. The extension of the model, accounting for workplace populations, demonstrated the extent of spatiotemporal fluctuation of working populations and demonstrated a marked improvement to overall revenue predictions.

Generally speaking, convenience stores still remain the worst performing predictions within the model. This problem is not unique to this research. Other academic studies have noted the difficulty in predicting convenience store revenue (Hood et al., 2015, Wood and Browne, 2007), noting that this may explain the lack of model-based applications and that often site visits and gut feeling have far more weight for retailers in the convenience market. However, many cluster catchments appear to demonstrate high levels of temporal fluctuation in demand, as demonstrated in chapter five, and thus have highly fluctuating revenue profiles, as observed in chapter four. Hence, all stores are impacted by temporal change and thus the addition of more temporal demand may go some way to resolve this problem. Over and under predictions in store revenue occur across the study region. This is especially felt for revenue predictions in the convenience market, where the temporal factors of trade, often within much smaller catchments than those found at supermarkets, have a greater emphasis on overall store revenue (Wood and Browne, 2007). This is perhaps the result of the poor level of representation of actual consumer behaviour and grocery demand, which may be improved through the increased level of temporal components. Understanding the temporal components within localised store catchments that make up the main trade areas of stores, which as demonstrated in chapter five can be highly variable, should improve the capacity of models. The increased insight on demand and the key trade drivers over time may partly offset the limitations of store revenue modelling techniques and weak performance of models, principally in the convenience market, as a result of the planned temporal extensions.

The following text indicates the perceived modelling components of each spatiotemporal demand type, such as expenditure and behaviour, which affect the interaction of consumers and retailers in the models. It is necessary to represent these consumer traits as optimally as possible in the SIM to ensure the extensions to the model are themselves robust. However, many of the aforementioned and fundamental behavioural traits incorporated into the SIM are experimental in design. In order to incorporate novel spatiotemporal components using newly available data this is unavoidable. Where individual components have a specific impact upon SIM operation, i.e. individual distance deterrence or expenditure estimates, extensions have been supported via corroborating evidence to limit the impact of researcher subjectivity.

8.2.1 Demand counts

Consumer demand counts, in terms of the volume of consumer populations and their locations throughout the day, for each distinct demand type mentioned, were discussed in chapter five. The spatiotemporal distributions are now implemented as novel demand layers in the SIM (see section 5.4 for more details).

8.2.2 Demand expenditure estimates and spatial allocation

In this section the demand expenditure estimates are discussed. These are novel developments that have rarely been attempted before in the literature.

8.2.2.1 University based expenditure

Data indicating the level of expenditure on groceries for university demand was generated from a number of different sources. Firstly, university household spend was derived via OACs which categorised students within supergroup two of the OAC: 'Cosmopolitans'. Subsequently, weekly household spend on food and alcohol was then gleaned from the Living Costs and Food Survey for the cosmopolitan OAC. Counts of student households are only a proxy, and they are derived from the average household size for student households in the UK for 2011: approximately 3.77 persons (Smith, 2014), and the total student population for each area. However, student properties are often purpose built and densely populated, with students favouring shared lets (Hubbard, 2009, Rugg et al., 2002). Whilst students' consumption patterns sometimes demonstrate that shopping activities are undertaken together, food is commonly purchased on an individual basis (Ness et al., 2002). Therefore, university expenditure estimates are based on the total student population for an area. This count is relatively easy to derive using the UK census, prescribing an individual weekly spend to each student. This individual weekly spend is reported to equate to approximately £20-£30 on average per student per week (Ness et al., 2002, ZenithOptimedia, 2016, Save the student, 2016). Based on the lifestyle evidence of student behaviour, which includes high consumption rates of alcohol compared to other typical consumers (Devine et al., 2006, Papadaki et al., 2007), the median threshold of £25 will be used for individual students. This accounts for average spending habits of students and typically results in higher total food expenditure per household than non-student households. Research found that students living away from home consumed fewer home-cooked meals (Papadaki et al., 2007). Similar research indicates that approximately 45% of students purchased food on or near campus multiple times during the week, (Pelletier and Laska, 2013). Following the same framework as workplace expenditure (discussed in chapter seven), a daytime expenditure of £3.25 per person per week was assigned to university students located on campus during the day. This was reallocated from the available student household spend. The daily allowance is based on the average price for common 'on-the-go' foods for 2016/2017, purchased by students close to Edinburgh university campus (International Office UOE, 2016). As noted in Pelletier and Laska (2013), this takes into account that not all students purchase food on campus.

8.2.2.2 School based expenditure

The estimated school based demand expenditure was based on the analysis of two recent consumer research surveys. The first survey reported the average pocket money received by children and the second survey reported teenage spending habits. Following the 2016 survey of teenagers, data suggests that teens spend approximately 20% of their available money on food (Piper Jaffyay & Co., 2016). The most recently available survey of national child pocket money (2015) suggests that the average child in the UK receives £6.20 per week (Mortimer et al., 2015). Following this logic, adolescents spend approximately £1.25 on average of their pocket money on food. Similarly, Caraher et al. (2014) reported that school pupils spent approximately £1.75 on snacks such as crisps, sweets and drinks during the school day, based on UK data from 2005. This value is higher than the expenditure estimate presented in this research. However, the reported differences are in line with the higher average pocket money figures in 2005: £8.37, which have since decreased to present day levels (Mortimer et al., 2015). Thus, the estimates appear logical.

8.2.2.3 Workplace expenditure

Workplace expenditure has previously been discussed in detail in section 7.2. However, to reiterate, workplace expenditure is set at a value of £5 per person per week for grocery spending.

8.2.2.4 Residential expenditure

Similarly to workplace expenditure, residential spending in the grocery sector has previously been well documented within this research. Expenditure estimates are based on the household OA super group classifications (ONS, 2011a) and the relevant average weekly household spend by each OAC, as reported in the Living Costs and Food Survey (ONS, 2015a). The methodology is discussed in section 6.2 above. However, for the extensions of the SIM incorporating temporal components, residential demand is modified to take into account temporal fluctuations of demand that result in a shift in the distribution of grocery expenditure over time. The movement of demand and shift in demand types in store catchments is represented through the inclusion of different demand types documented above. Weekly grocery expenditure is reallocated to correlate with the temporal shifts resulting in the daytime, redistributing the allocated expenditure from the total grocery spend at the location of the household to the daytime location of the individual. For instance, workplace expenditure is set at £5: then if ten workers live in an area and work somewhere else, a total of £50 is removed from the available expenditure in that area (in this case the LSOA) and redistributed to the relevant WPZ. Similarly, the total number of student households in any one area is subtracted from the total residential count of households using their original OAC average weekly food spend, to maximise the accuracy of residential expenditure estimates.

8.2.2.5 Leisure expenditure

Based on evidence adapted from Orrell et al. (2015) in table 5.3 above, the approximate total weekly expenditure for individual visitors in West Yorkshire is £29. The national proportion of visitor expenditure for *Eating and Drinking: Food bought in shops/takeaways & consumed on trip* in 2015, source: Table 3.9 - p78 Orrell et al. (2015), was used to derive visitor expenditure on food purchased in stores. Spending in this category was observed to account for 5.4% of the total expenditure for tourists. Hence, an expenditure estimate of £1.56 on 'groceries' is applied per person per week to visitors. Visitor expenditure is important to account for instore revenue predictions because although the relative expenditure values per person are small, the contribution from the aggregate spending power by visitors, particularly at certain times of the year and in certain locations, will be considerable (as demonstrated by Newing et al., 2013). Day visits to attractions only makeup one segment of total leisure demand. However, accurate spatially referenced data on total visitor numbers is difficult to obtain. Therefore at this stage, any additional expenditure generated by other visitors, such as additional numbers driven by visits to friends and family, is excluded from this research.

8.2.3 Expenditure outputs

Presented below are a series of outputs summarising the expenditure estimates in West Yorkshire for each of the demand groups being used.

Table 8.2 - Total expenditure estimates for demand groups in West Yorkshire

Demand group	Expenditure estimate (£ - Millions)
Traditional household	61.46
Residential	54.24
University student	2.47
Campus based	0.27
School based	0.19
Leisure (visitors at attractions)	0.12
Workplace	5.18
Daytime total	62.47
Daytime Difference	1.02

Figure 8.1 – Map showing traditional household grocery expenditure - distribution based on household OAC average weekly spend reported in the Living Costs and Food Survey

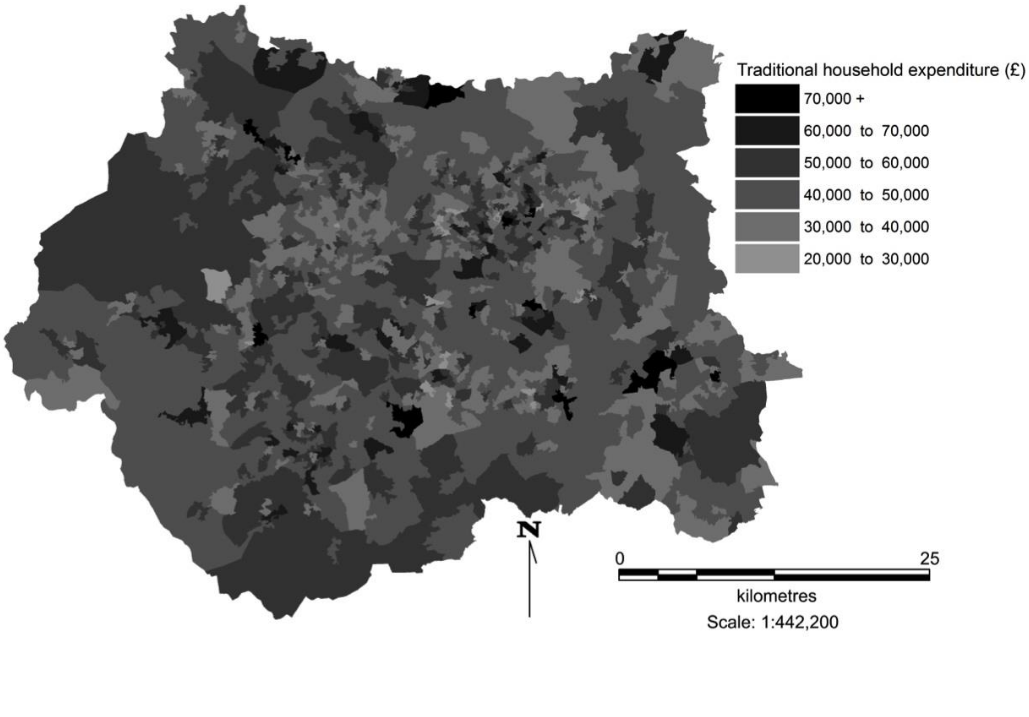


Figure 8.2 – Map showing residential expenditure following the reallocation of expenditure to different spatiotemporal demand types

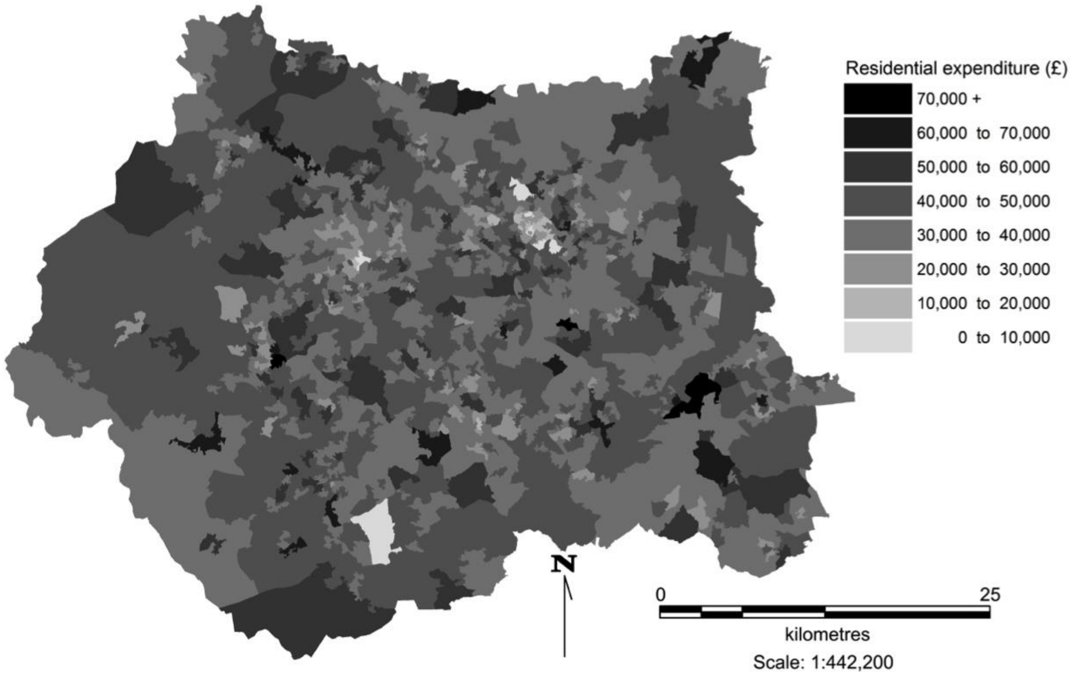


Figure 8.3 – Map showing university demand grocery expenditure according to usual place of residence (this accounts for the reduction in expenditure resulting from the reallocation of expenditure spent on student campuses during the day)

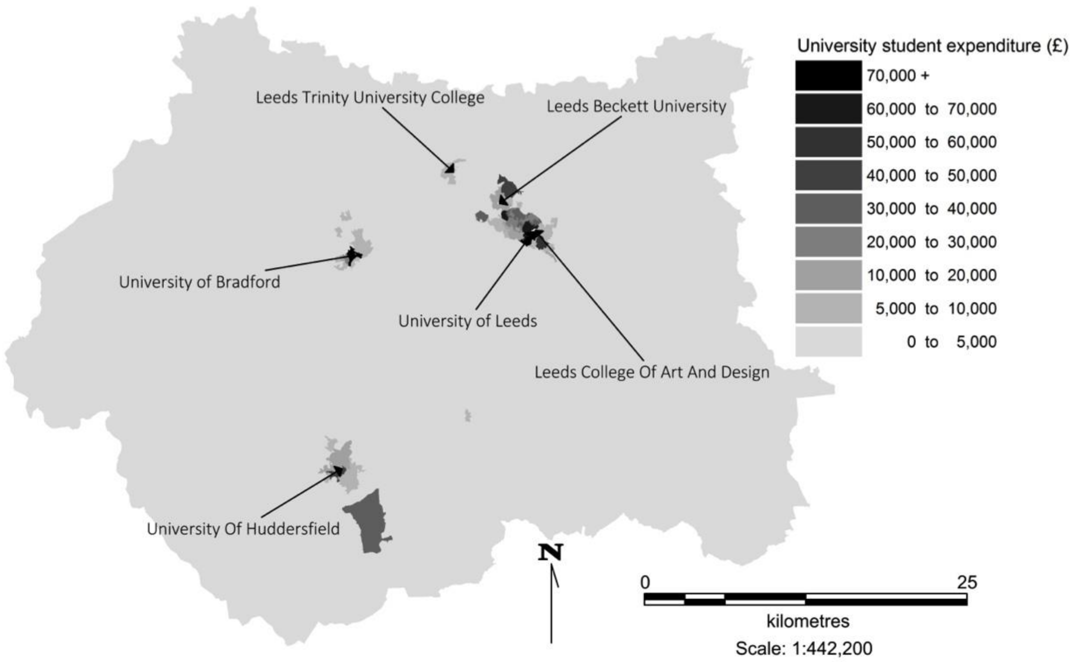


Figure 8.4 – Map showing University Campus based student grocery expenditure according to approximate midday population estimates

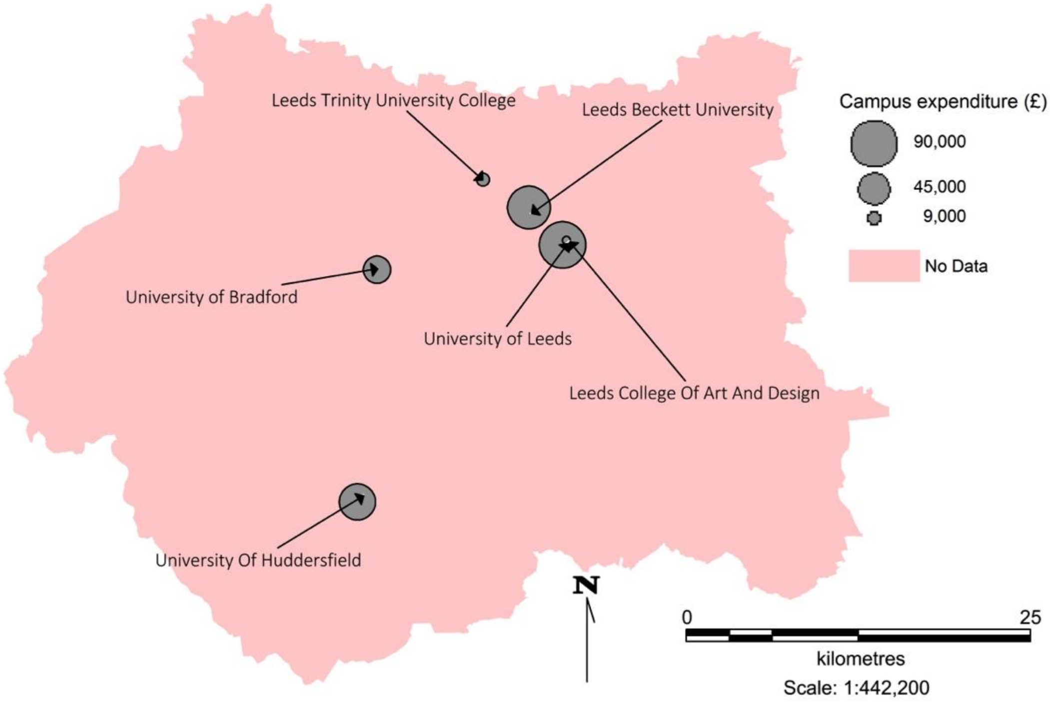


Figure 8.5 - Map showing secondary school based expenditure during the school day

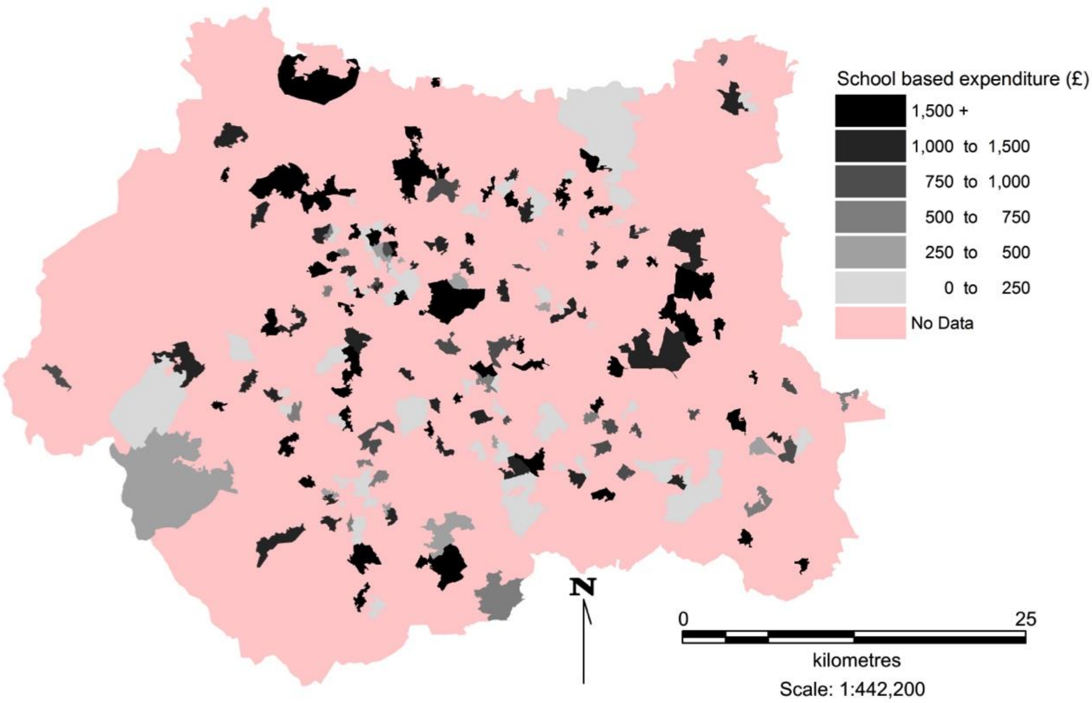


Figure 8.6 - Map showing leisure expenditure of potential store bought food by main attraction locations

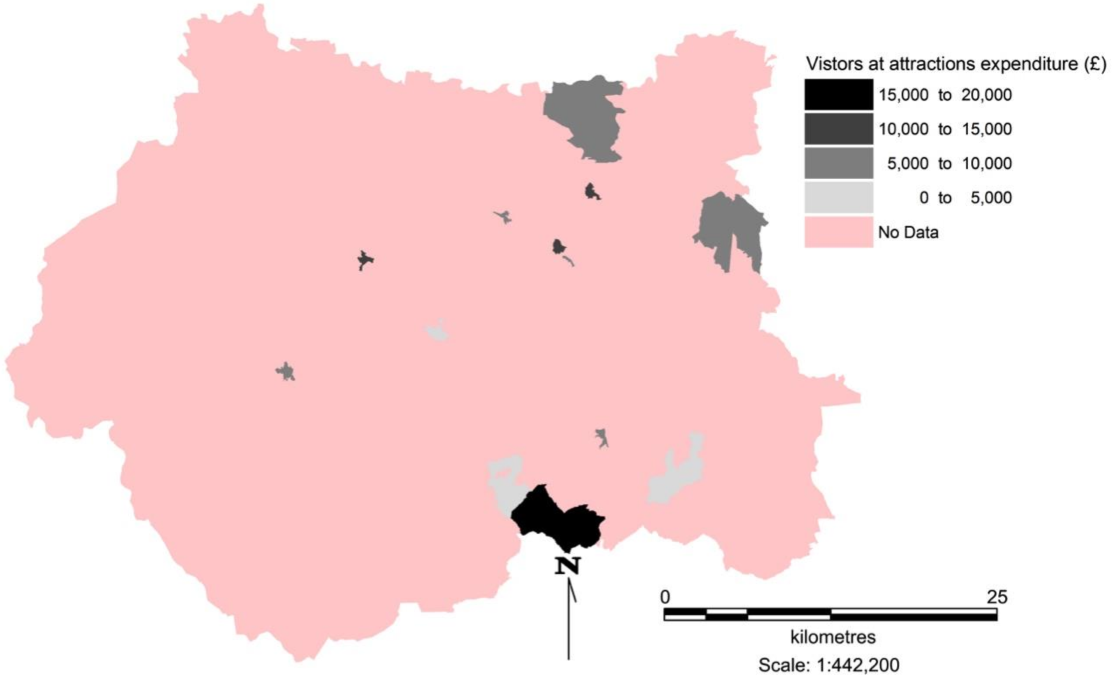
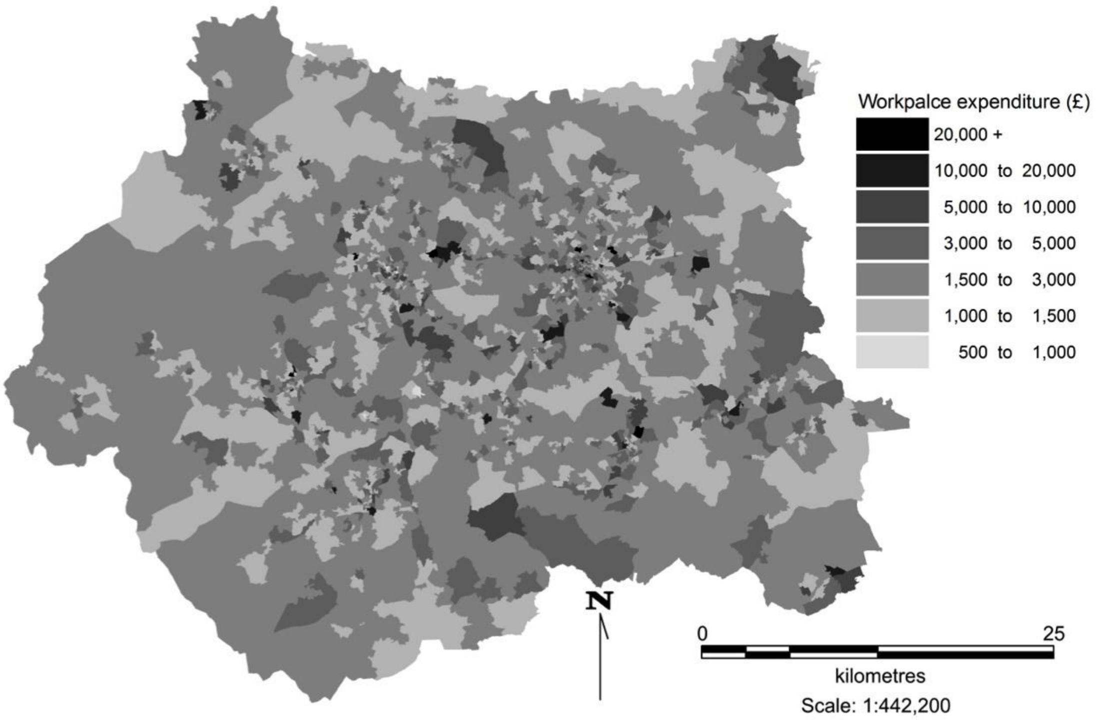


Figure 8.7 – Map showing workplace expenditure by workers usual place of work



8.2.4 Spatiotemporal behaviour

In this section the central components of each demand group's grocery shopping behaviours are discussed. It is important to account for the inherent traits of each demand type, as differences in behaviour and shopping attitudes will result in a varied interaction with retailers throughout the day. This is likely to have a subsequent effect on store revenue predictions, especially if major fluctuations in demand also occur over time.

8.2.4.1 University demand behaviour

A recent survey of student grocery shopping behaviour found that 98% of food shopping was done in grocery retailers, with 70% in supermarket brand stores and the remaining 28% in symbols and independents. Furthermore, it appeared that 62% of students shopped for their groceries within 1 mile of their accommodation, with a further 22% (84% in total) shopping within 2 miles (Devine et al., 2006). It appears that students will favour stores close to their home and that they prefer to shop at major retail brands. Consequently, store attractiveness (weighted by alpha in the model) will be higher for major retailers to account for this behaviour. Similarly, 60% of students were observed to walk to grocery stores (Devine et al., 2006). Hence, the distance deterrent factor will be higher than for typical residential behaviour to account for the tendency to shop locally and the reduced mobility of students. This appears to correlate with the drop off distances for actual sales observed in cluster four stores, which have catchments populated by large student cohorts and after a short distance, the observed proportion of store sales drops markedly, as shown in Table 5.2. The second component of university demand is the daytime campus population. Student behaviour here is likely to follow the same store preference tendencies, i.e. favouring major grocery retailers. However, in this instance the distance deterrent will be much higher. Analysis of loyalty card data and store revenue data for cluster four stores appears to demonstrate very little impact on revenue for stores over 500 meters away from the presence of a sizeably university campus during the day (see figure 4.11 and Table 5.5). It appears that students who purchase food on campus are highly sensitive to distance and appear to avoid longer distances in the process of purchasing food. As a result, a considerably higher distance deterrent values will be used to reduce the distances travelled by 'on-campus demand' to ensure expenditure remains located within the campus locality.

8.2.4.2 School based demand behaviour

Evidence suggests that school students are relatively immobile and that they appear to operate within relatively small catchments surrounding their school when consuming food and snacks (Caraher et al., 2014, Cavill and Rutter, 2013). Academic and government backed planning policies suggest that school pupils are unlikely to travel further than a maximum distance of

800m (a ten minute walk) and appear to more commonly stay within a 400 meter buffer distance of the school, equating to a five minute walk, when purchasing food (Caraher et al., 2014, Cavill and Rutter, 2013). In the case of the government and council led initiatives reported by Cavill and Rutter (2013), these distances were used as planning guidelines for exclusion zones of fast food outlets surrounding schools. There has been little attention on the wider food environment surrounding schools (School Food Trust, (2008) cited in Caraher et al. (2014) p.1), hence the planning policies undertaken by the government regarding the distances covered by school pupils in the consumption of food will be used for this research as other evidence is limited. This results in school based demand receiving a high beta value to account for the effect of distance as a major deterrent. Subsequently, a maximum threshold was set in the model, restricting flows to stores over a distance of 800 meters, in line with current government guidelines. The purchasing habits of school pupils are focused on the purchasing of snacks such as crisps, drinks and sweets. Furthermore, aside from the purchasing of food from takeaways and fast food outlets, the propensity for students was to purchase these items in local shops (Caraher et al., 2014, Cavill and Rutter, 2013). Fast food and takeaway consumption are not the focus of this study and as such this aspect of consumer behaviour is excluded in the model: this is accounted for in the expenditure estimates balancing daily and weekly demand so that only money spent in grocery stores is represented in the model. Based on the typical food and drink products bought, school students are unlikely to favour any grocery retail brand over another. As major supermarket stores and local independents, for example, typically all stock the range of the snack based products primarily purchased by school pupils, the subsequent store attractiveness for school based demand will be neutral for all brands, making distance the primary factor affecting store choice.

8.2.4.3 Workplace demand behaviour

Workplace demand was discussed in detail in the previous chapter and hence, this will not be discussed in detail again. In short, a consumer survey on workplace grocery consumers suggested that workers are highly time dependant and consequently they are sensitive to the costs incurred by travelling to a store, i.e. the distance or time it takes to get there. The survey found that workplace based consumers typically choose stores based on convenience and that generally they would purchase food within 500 meters of their place of work. Subsequently, a high distance deterrence value is used in the SIM to replicate workplace trips. However, unlike school-based demand for instance, no maximum threshold is set which would create a cut off distance. This is because it is likely that workplace demand is more mobile. In some instances consumers will have access to a car, and in more rural areas consumers will likely travel further to purchase their food while to and from, or at work. Therefore, while localised stores receive higher priority, workplace flows may still cover greater distances when presented with limited food purchasing options nearby.

8.2.4.4 Residential demand behaviour

Residential demand was previously discussed in chapter six. Demand side behaviour will remain the same, with flows originating from the residential household to stores following operation as before. For instance, travel behaviour and store preference are disaggregated by household type and have been suitably calibrated using observed consumer data. The major difference to residential demand is the reduction of available expenditure during the day, which accounts for the behaviour of consumers as they go about their daily activities and monies are reassigned to the different demand groups noted above.

8.2.4.5 Leisure demand behaviour

The theoretical assumption regarding leisure behaviour, in this instance the behaviour of day visitors to specific tourist attractions, is that visitors are likely to concentrate their time and energy in, or within a close proximity to, the attraction destination (Lew and McKercher, 2006, Shoval and Isaacson, 2009). The attraction is typically the main reason for visiting and as a result beta values are set relatively high to reduce the likelihood of unrealistic journey times to stores further away. An experimental optimal value of around 500 meters was used for the typical ATD, as individuals are likely to be on foot once at a destination. Furthermore, in some instances they may be unfamiliar with their surroundings and consequently they are unlikely to venture far. Similar to school based demand, a maximum threshold on the distance flows can originate from was also applied. This is to avoid adding in additional expenditure in situations where the presence of a daytime visitor population is unlikely to have an impact on grocery store revenue. This situation may arise when the attraction is rural and no local grocery stores exist: thus spending is unlikely to happen. A threshold was adopted to stop the SIM from falsely assigning expenditure to the nearest stores when in reality the likelihood that flows will occur is minimal.

8.3 Daytime spatial interaction model: spatiotemporal extension

8.3.1 New SIM equation

Following the disaggregation of spatiotemporal demand types, the new SIM equation, which incorporates the above mentioned expenditure and behavioural tendencies of each demand type, is expressed below. The model is split across the two geographies (LSOA and WPZs) and combines the store sales outputs from all demand types to produce a final store sales prediction.

Individual spatiotemporal demand is expressed as follows: $h \in [h_1, h_2, h_3, h_4, h_5]$, where:

h_1 is the total university student demand from residences h_2 is the total available campus based university student demand, h_3 is the total school based demand, h_4 is the total leisure demand and h_5 is the total daytime residential demand by household type g.

$$S_{ij}^{hbf} + S_{kj}^{bf} = \sum_h [(A_i^h O_i^h W_j^{\alpha^{hbf}} \exp^{-\beta^h c_{ij}})] + (A_k O_k^z W_j^{\alpha bf} \exp^{-\beta C_{kj}})$$

(8.1)

where:

S_{ij}^{hbf} represents predicted expenditure between origin i and store j by demand type h and store brand b and store format f. Store format f differentiates between store formats based on whether store j is a supermarket or a convenience store.

S_{kj}^{bf} represents predicted expenditure between workplace zone k and store j by brand b and store format f (equation *(7.1)* defined in section 7.4).

O_i^h refers to total grocery expenditure for demand type h in origin zone i.

O_k^z represents the total grocery expenditure in workplace zone k.

$W_j^{\alpha^{hbf}}$ is the measure of attractiveness of store j to the power of α^{hbf} $where$ α^{hbf} is a power function influencing the importance of the attractiveness parameter for store j by demand type h and store brand b and store format f.

$exp^{(-\beta^h c_{ij})}$ a distance deterrence factor, incorporating changes in behaviour for demand type h, impacting the distance travelled between origin i and retail destination j.

A_k as expressed above in equation 7.3

A_i^h represents the balancing factor accounting for competition in the model and ensuring that all demand from origin i by demand type h is distributed to stores within the study area. It is calculated as:

$$A_i^h = \sum_b \frac{1}{\sum_j W_j^{\alpha^{hb}} exp^{(-\beta c_{ij})}}$$

(8.2)

8.3.2 Daytime SIM analysis

The current model performance in relation of the predicted store sales compared to actual observed revenue (as a percentage) for the daytime model (Table 8.3) is presented below. The daytime model is the refined time-of-day fit using spatiotemporally disaggregated demand. The distribution of this demand has previously been demonstrated in chapter five, see Figure 5.19.

The store predictions are presented alongside the night time model for comparison, to assess how the model predictions have evolved following temporal refinement. The night time model is the disaggregated SIM built in chapter six, which represents a residentially based population only (typically representative of a night time population). The spatial distribution of the night time model demand is likewise discussed in chapter five. Overall, there are an increased number of stores with a prediction within an acceptable performance threshold, suggesting an improvement in the capacity of the model. The average store prediction for all stores is 90% of observed store sales. For supermarkets, the average store prediction is 103% and convenience stores 83%, which is an increase from the previous night time only model (see Table 7.2).

There are a total of eight stores now with a prediction within the +/-10% threshold compared with only two in the original model. Table 8.4 demonstrates the magnitude of the changes in accuracy of store predictions, through the percentage of change grouping stores into three categories: those under-predicted, over-predicted and those within the 10% threshold. There has been a decrease in the number of stores under-predicted (although there are still several stores that under-predict) and an increase in the number of stores within the acceptable accuracy threshold of +/-10%. However, there has also been a sharp increase in the number of stores now over-predicting revenue.

Table 8.3 - SIM accuracy of predicted store sales as a percentage of observed sales

	Night time Model			Daytime Model		
% of Actual Revenue	All stores	Supermarket	Convenience stores	All stores	Supermarket	Convenience stores
Count <50	20	1	19	9	0	9
Count 50-60	7	4	3	2	0	2
Count 60-70	7	1	6	5	2	3
Count 70-80	6	6	0	4	0	4
Count 80-90	4	2	2	5	3	2
Count 90-100	2	2	0	5	2	3
Count 100-110	0	0	0	3	2	1
Count 110-120	2	0	2	4	3	1
Count 120-130	0	0	0	4	2	2
Count 130-140	0	0	0	2	1	1
Count 140-150	0	0	0	1	1	0
Count > 150	0	0	0	4	0	4

Table 8.4 - Percentage of change of store predictions between the night time and daytime model

	% change in store predictions		
	All Stores	Supermarkets	Convenience
Within +/-10% threshold	300.00%	100.00%	4 additional stores
Below +/-10%	-43.18%	-64.29%	-33.33%
Above +/-10%	650.00%	7 additional stores	300.00%

To illustrate the overall improvements to realism in store predictions and the differences between the night time model and the daytime model, figures 8.8 and 8.9 demonstrate the flows (sales in £) to a city centre store within its localised catchment area. Incorporation of spatiotemporal components of demand results in an increased level of spatial insight into demand type spending and temporal availability. This particular example demonstrates an increased volume of total flows to the store within both the 5km and 1km buffer (Table 8.5). This is likely far more representative of the conditions at this store expected in reality. A notable difference too the daytime model is the considerable volume of flows originating from workplace demand and which represent a substantial proportion of total flows to the store, sales that were restricted before when only using a residential population in the night time model. Furthermore, upon observation the spatial pattern of flows and the individual behaviours of different demand groups appear to be operating as expected, i.e. distances travelled are within typically reported behavioural patterns, indicating a more realistic platform for modelling. To assess the model reliability and whether any bias exists, a statistical analysis on store sales was undertaken reviewing the distribution of store predictions.

Figure 8.8 - Flows (£) to city centre store based on the night time model (each dot represents £50)

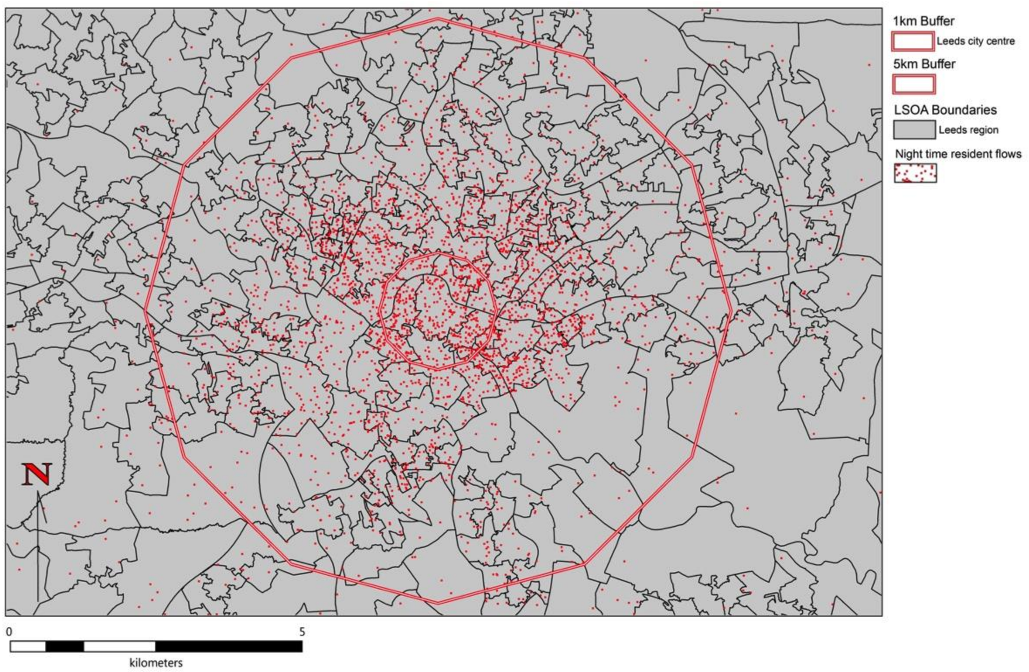


Figure 8.9 - Flows (£) to city centre store based on the daytime model (each dot represents £50)

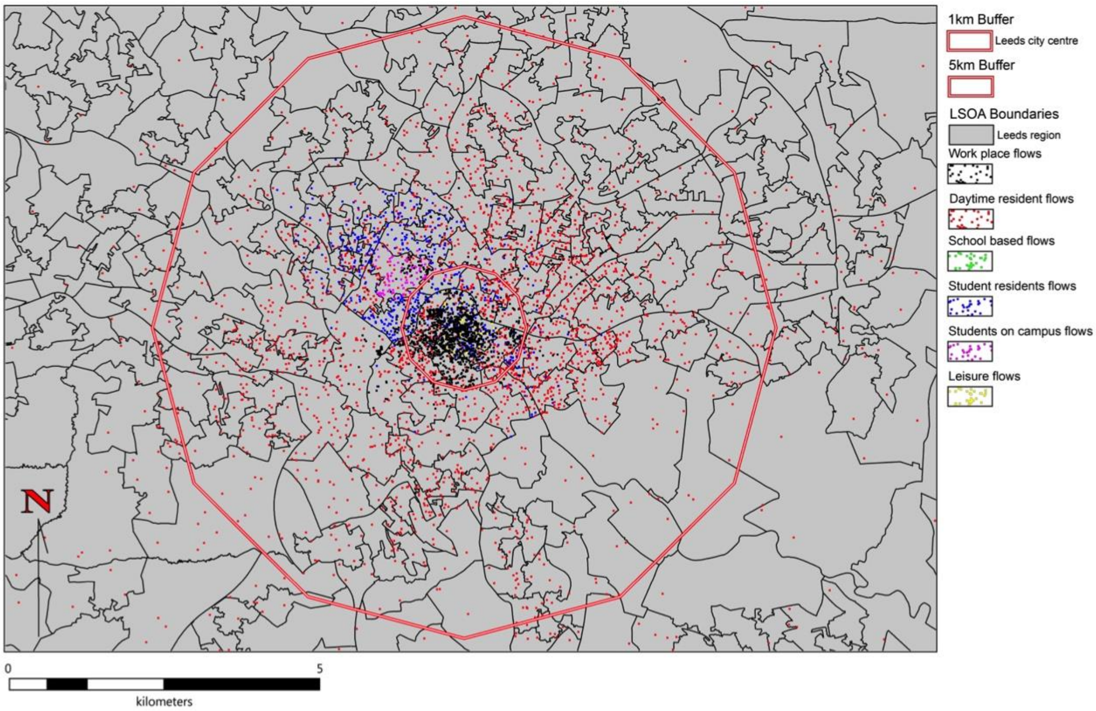


Table 8.5 - Summary of flows to a city centre store for night time and daytime model

	Total flows to store (£)	
	Night time model	Daytime model
5km buffer	£161,277.00	£235,202.00
1km buffer	£22,949.00	£82,386.00
Predicted store sales accuracy	84.00%	115.00%

Using the boxplot presented below (Figure 8.10), the model was observed to have a normal distribution but still demonstrated a tendency to under-predict store sales for convenience stores and to over-predict store sales for supermarkets: this was observed in the city centre example store (Table 8.5). This suggests that there may still be potential limitations in store attractiveness within the model that any future research should consider addressing, but could also be a result of limitations in the general SIM approach. The current standard deviations in predictive accuracy for all store types, particularly for c-stores, generally fall outside of a 10+/- threshold (shown below in figures 8.11, 8.12, 8.13). The variance observed in the accuracy of the model suggests that additional aspects of consumer behaviour lost through the aggregation of consumer types could still be missing from the model. Consumer behaviour is far more complex, resulting in the localised variation in predictive accuracy seen for some stores. Hence, the current accuracy, is an improvement on the previous model. Nevertheless, the impact and importance of localised spatiotemporal demand are apparent and at least demonstrate spatially and temporally fluctuating variations in demand. Needless to say, additional behaviour traits effecting grocery spend (as discussed in chapter two) which may represent influential components important for future SIM development. For example, the partner's stores located in Harehills and Bradford represent prime examples of this issue. Following the incorporation of spatiotemporal demand and supply and demand disaggregation, which to the best of our ability are robust, store predictions remain considerably higher than those observed. Interestingly, these areas represent considerable proportions of non-white ethnic populations, which as noted in chapter two, demonstrate very different attitudes in grocery shopping behaviour (when all other things remain equal) to other consumers. This may explain the model's tendency to over-predict stores in this type of location whereas in reality a much higher volume of flows could go to local independents and a more up to date dataset including all ethnic independents would likely resolve this issue.

Figure 8.10 - Boxplot of predicted store sales accuracy

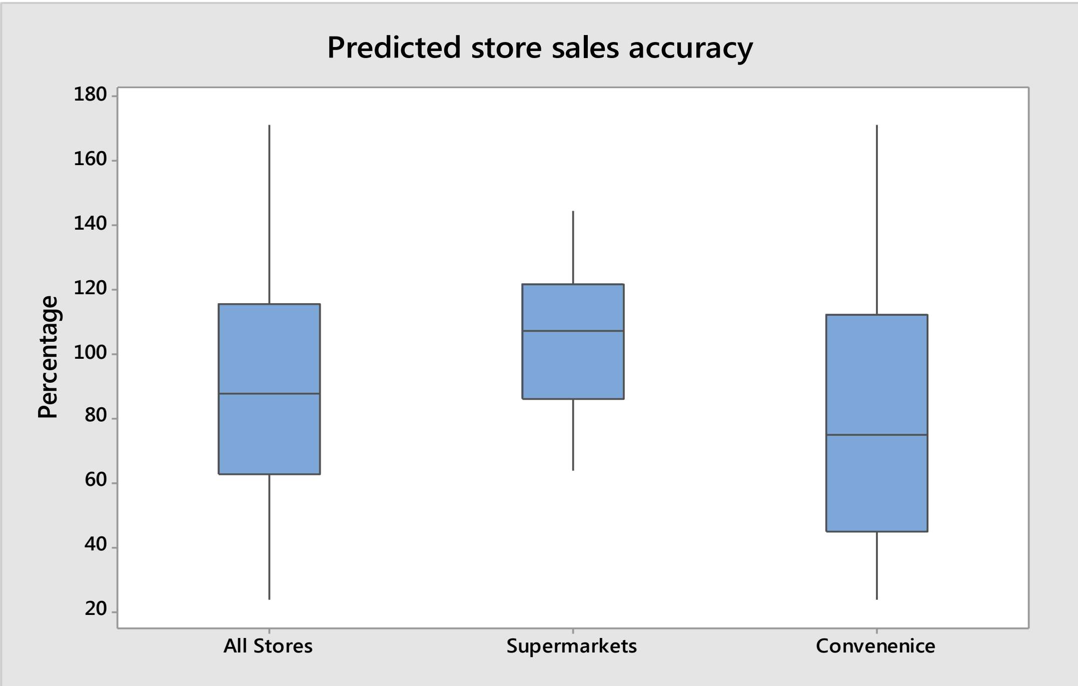


Figure 8.11 - Statistical summary of daytime model store predictions (for all stores)

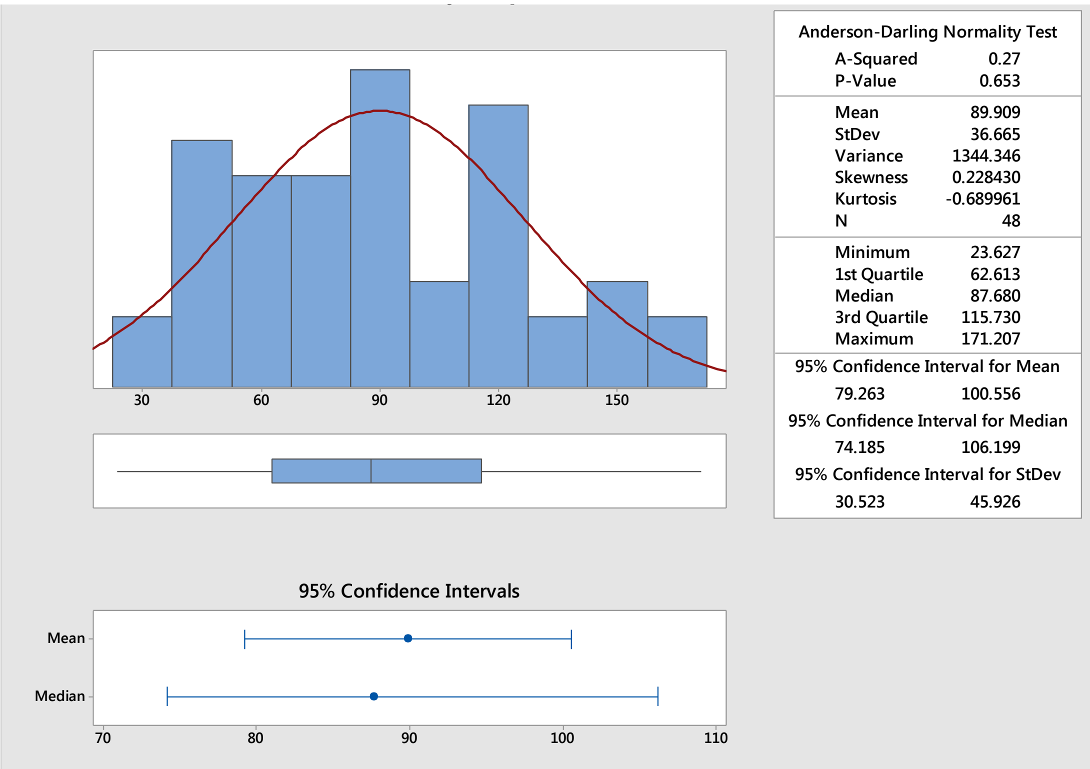


Figure 8.12 - Statistical summary of daytime model store predictions (for supermarkets)

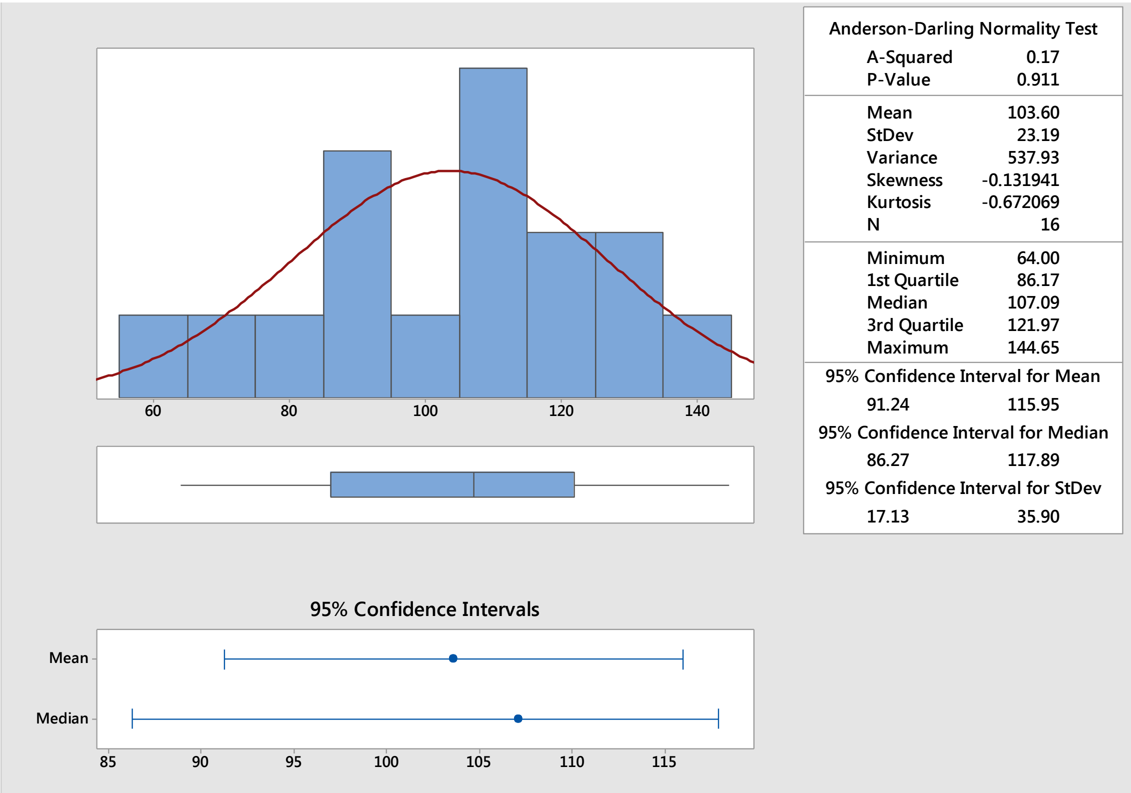
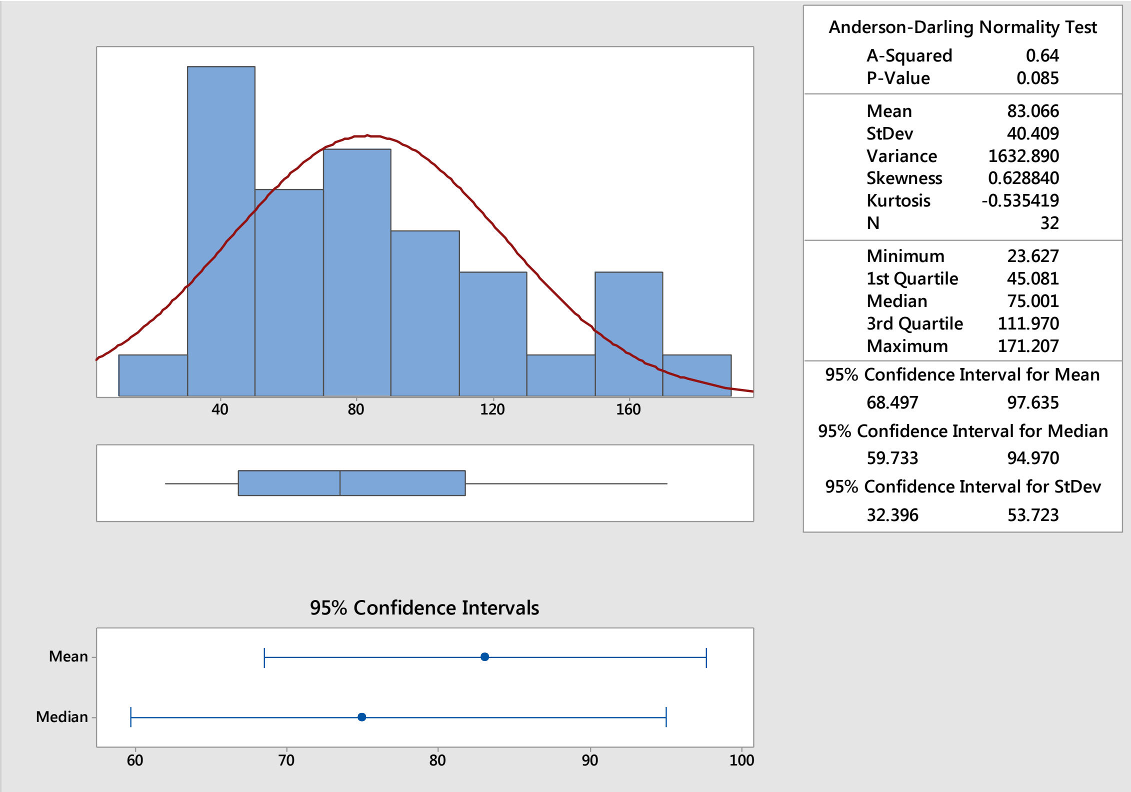


Figure 8.13 - Statistical summary of daytime model store predictions (for convenience stores)



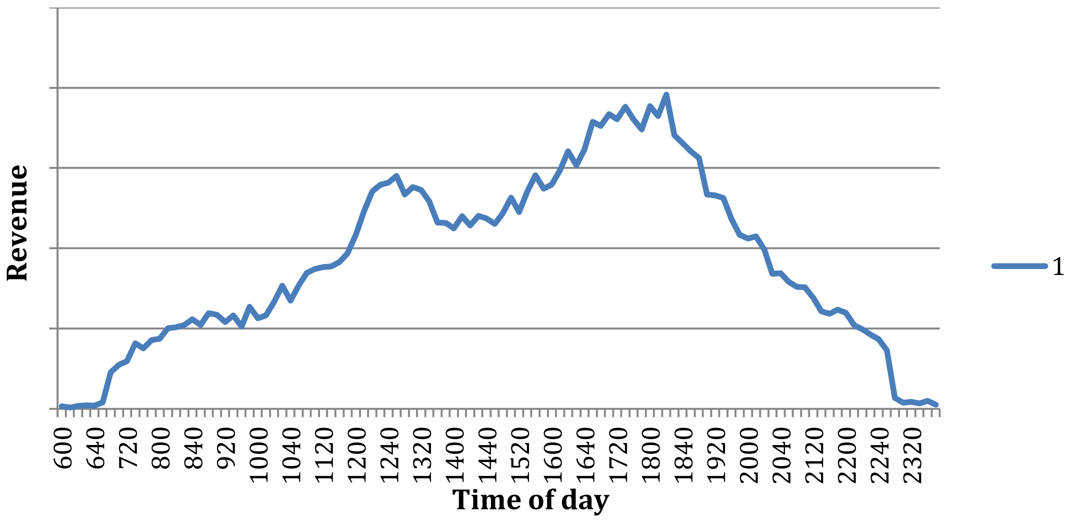
Much of the novel demand developments have been experimental, utilising, not only new datasets, but also those typically not accessible in academia due to their commercial nature. Where possible, grocery behaviour (especially those situated within the novel demand groups) was supported with relevant literature as evidence. While there are potential drawbacks to this approach, i.e. such as consumer behaviours not being validated with observed data; the current lack of attention on spatiotemporal demand types in the field of grocery store modelling and limited availability of data, the calibration of the model with theoretical evidence from literature is likely to be at present the most feasible option within the scope of this study. Demand was calibrated in terms of brand preference, expenditure and propensity to travel. For instance, the propensity to travel was adjusted for school-based demand until the ATD fell within the stipulated behavioural attitudes. Once these conditions were replicated, it was assumed that optimal conditions were achieved and any further manipulation of the data would be subjective and thus could result in user bias. In section 8.4 discussion and evaluation of the current model's representation of typical demand behaviours and the typical supply and demand-side observed behaviour are compared.

8.4 Daytime model: Model behaviour VS reality

It is necessary to undertake a form of quality control comparing the model with 'expected' norms. This is important because, as mentioned, many of the spatiotemporal concepts discussed in this research are novel with regard to store location planning. Therefore, comparisons of the typically expected behaviours are essential. This not only checks model quality, but also affords useful insight into aspects of grocery retail and consumer behaviour via the incorporation of spatiotemporal data, as well as any potential weaknesses of the model. Not only will this be beneficial to this research, but it will also offer novel insight for retailers with potential implications for modelling and operational applications with benefits at a store level.

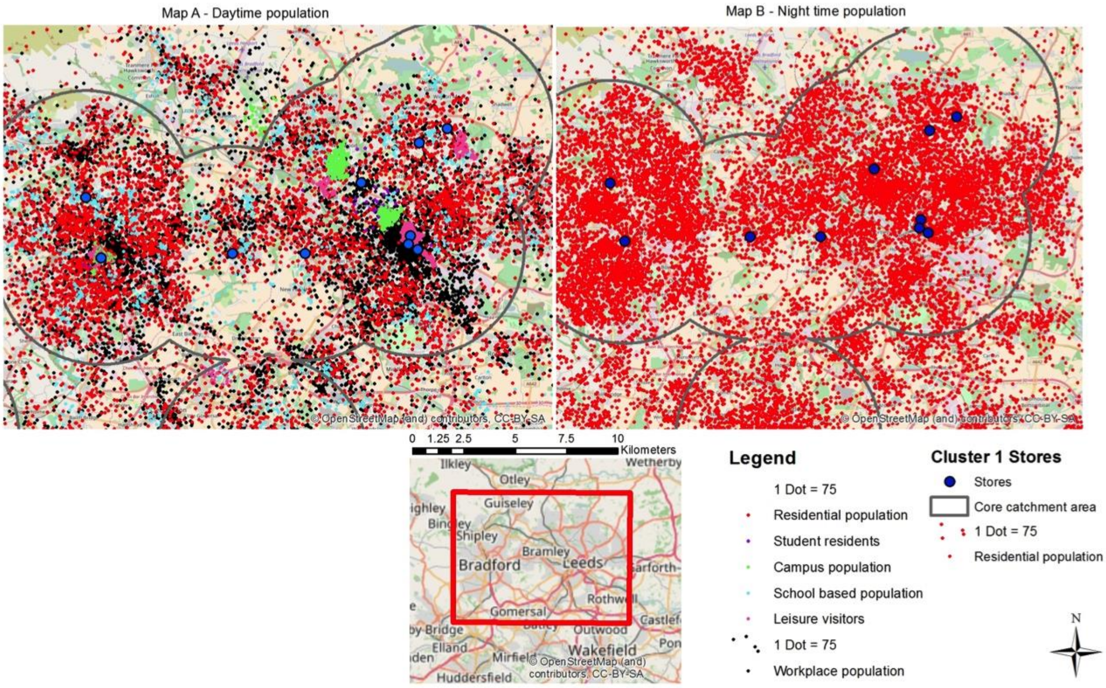
Cluster one stores - '*workplace convenience'*:

Figure 8.14 - Temporal sales profile for cluster one stores, averaged over a week



Stores in cluster one were classified as 'workplace convenience', appearing to demonstrate a relationship between workplace demand during the day as indicated via the peaks in sales at certain points during the day. Stores were predominantly convenience stores as well as being located in or near urban centres. Analysis of the core catchments using an aggregate residential night time population and a disaggregate population by day was undertaken early in chapter five. This revealed that there was a stark contrast between the spatial distributions of a census based residential populations (typical of night time population distributions) and the spatial pattern of a disaggregated daytime population. Ultimately, not only does a residential based population fail to represent diurnal spatiotemporal fluctuations in the population, but it also fails to accurately represent the varied consumer dynamics of individual segments of the population. In the core catchment area for this cluster, there was a net increase in the total population during the day, indicating that store sales should peak during the day. Additionally, there is a considerable shift in the makeup of this population, not only in terms of numbers, but also in consumer behaviour and the spatial distribution of the population, switching from a residentially dominated population at night, to a workplace dominant population during the day. This is illustrated in Figure 8.15 for Leeds and Bradford. A residual, yet lower residential population appears to remain present during the daytime too. This spatial pattern is particularly apparent within close proximity to the stores, which typically went from low to high numbers especially in regard to the number of workers, hence the cluster group name. Similarly, there is also a sizeable campus based population during the day within the core catchment area for this cluster, although their behaviour is highly distance dependant so the impacts may be minimal.

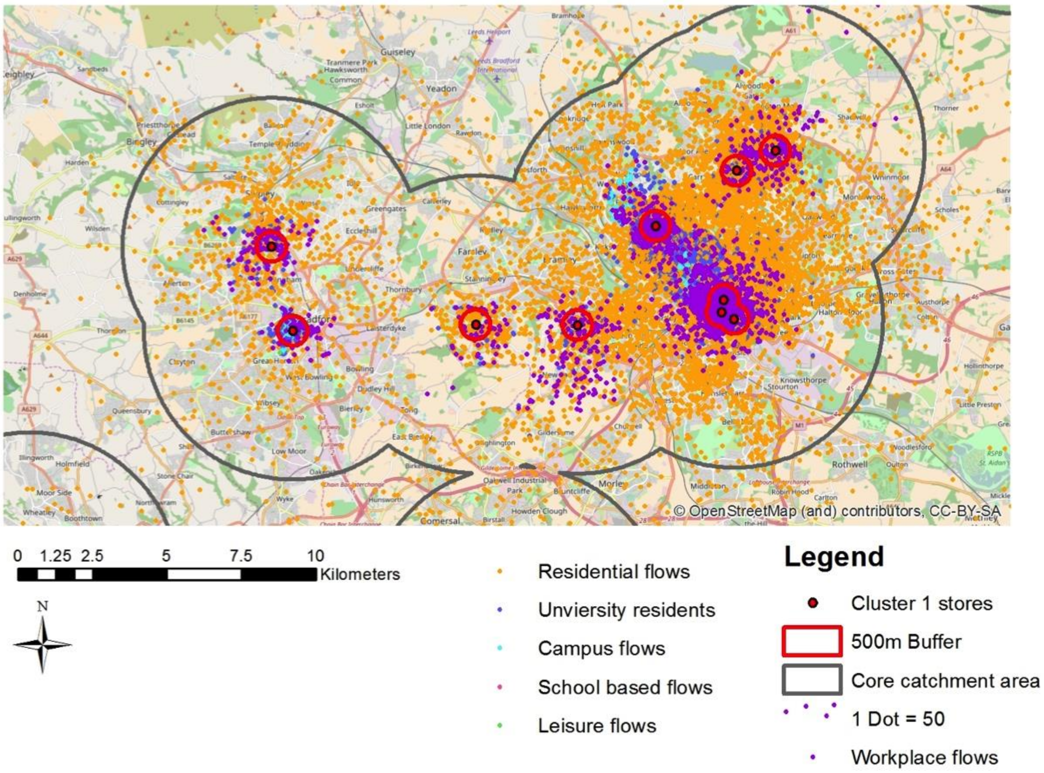
Figure 8.15 - Diurnal population distribution within cluster one core catchment areas, Leeds and Bradford (each dot represents 75 people)



Based on the above observations, data analyses and the expected behaviours that are typically associated with the demand types represented in cluster one, one would expect to observe a series of distinct spatiotemporal and behaviour patterns. For instance, store sales vary considerably during the day and increases in sales are likely to correspond with the arrival of workers (Berry et al., 2016, Schwanen, 2004). Particularly busy periods of trade will result from lunchtime sales and end of day sales, which will then decline as daytime based populations leave the area (as demonstrated in figure 8.14). Sales at these stores will originate from locations in close proximity to the store due to the characteristically shorter distances, typically travelled by the demand types present in the cluster one core catchment (Berry et al., 2016, Devine et al., 2006). A sizable proportion of the store revenue will probably be driven by the nearby workforce population.

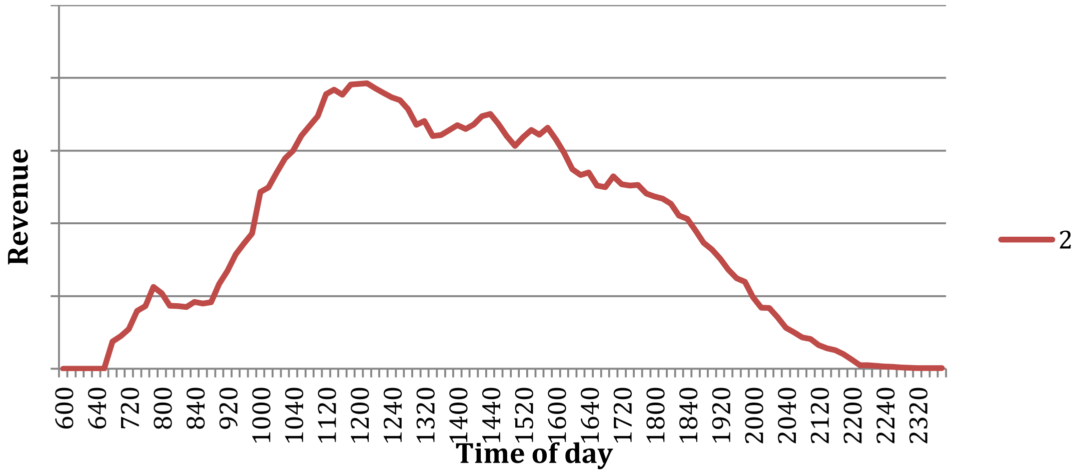
Over 90% of the flows to cluster one stores originate from within the core catchment area (5km) in the model. The reality that core catchments derived from loyalty card sales may be too extensive; actual sales that originate from further away (at the home address of the registered loyalty card user) are in reality the result of sales occurring from within close proximity to the store, such as the workplace (as noted in chapter five) and is observed in Figure 8.16. The map showing flows to cluster ones stores in Bradford and Leeds illustrates how in the model the majority of flows originate from much closer to the store. While this illustrates the drawback of deriving core catchments from loyalty cards, as noted in chapter five, as well as demonstrating potential indicators for spatiotemporal demand in the loyalty card data, this behaviour (i.e. flows clustered around stores for these particular demand types) is a better representation of the expected norms, particularly for workplace dominated stores and the workplace consumers. This may explain the high volume of flows in the city centre originating from residential flows from the surrounding suburbs. Needless to say, it was necessary to limit the propensity to travel for all demand types further. The distance travelled by consumers appears too high, e.g. University demand and workplace demand are travelling in excess of the typically expected norms as indicated in the map below, travelling over 500m. Nevertheless, in the cases of cluster one it appears that the model handles the differences in behaviour of each demand type well, there are clear distinctions between the behaviours of each demand type and the temporally dependent distributions appear realistic.

Figure 8.16 - Flows (£50 per dot) to cluster one stores by demand group, Leeds and Bradford



<u>Cluster two stores '*Traditional supermarkets*':</u>

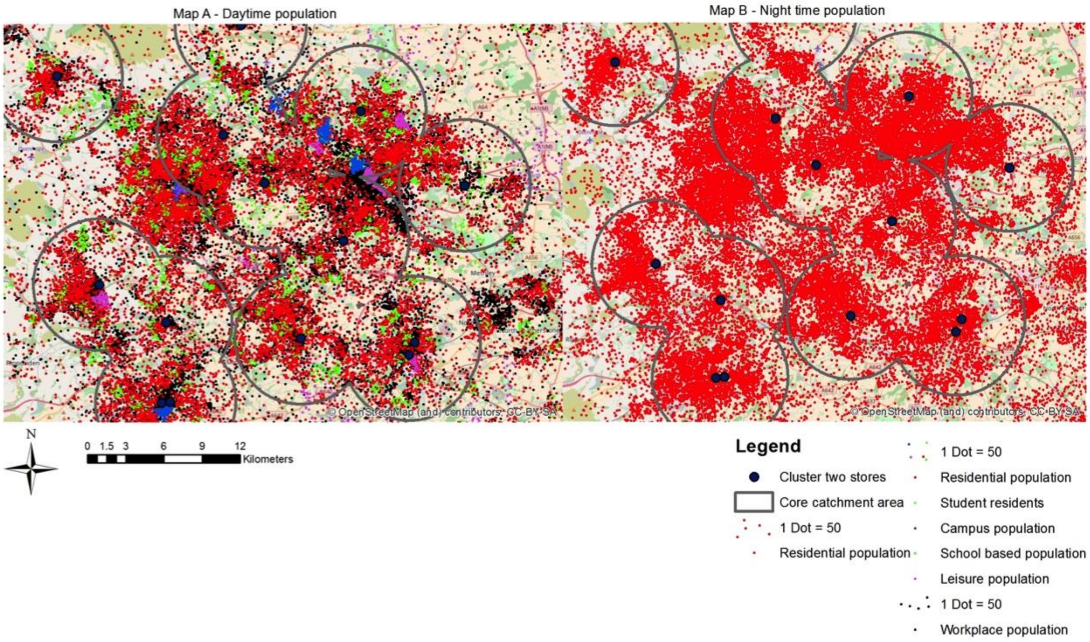
Figure 8.17 - Temporal sales profile for cluster two stores, averaged over a week



Stores in cluster two are primarily the supermarket format and have the largest average floorspace of all the cluster groups. The sales profile of these stores (Figure 8.17) follows the characteristic diurnal pattern associated with traditional supermarket grocery shopping behaviour throughout the day, seemingly unchanged for the last two decades (East et al., 1994, Ipsos-RSL, 2003). Interestingly, the implication of a consistent and 'traditional' sales profile suggests that the behaviour of consumers shopping at these stores has remained relatively constant over the years. Many of the stores are located at out of town sites or near other shopping facilities and retail parks. The population within the core catchment of cluster two is predominantly residential, which remains high during the day (Figure 8.18). The catchment is also populated by high and spatially clustered workplace populations at several of the sites within the cluster. Consequently, the biggest temporal shifts occur in close proximity to the store. The high residential night time population and high volume of day time residents, in addition to the high volume of workers employed in the surrounding businesses means there is a considerable volume of available demand.

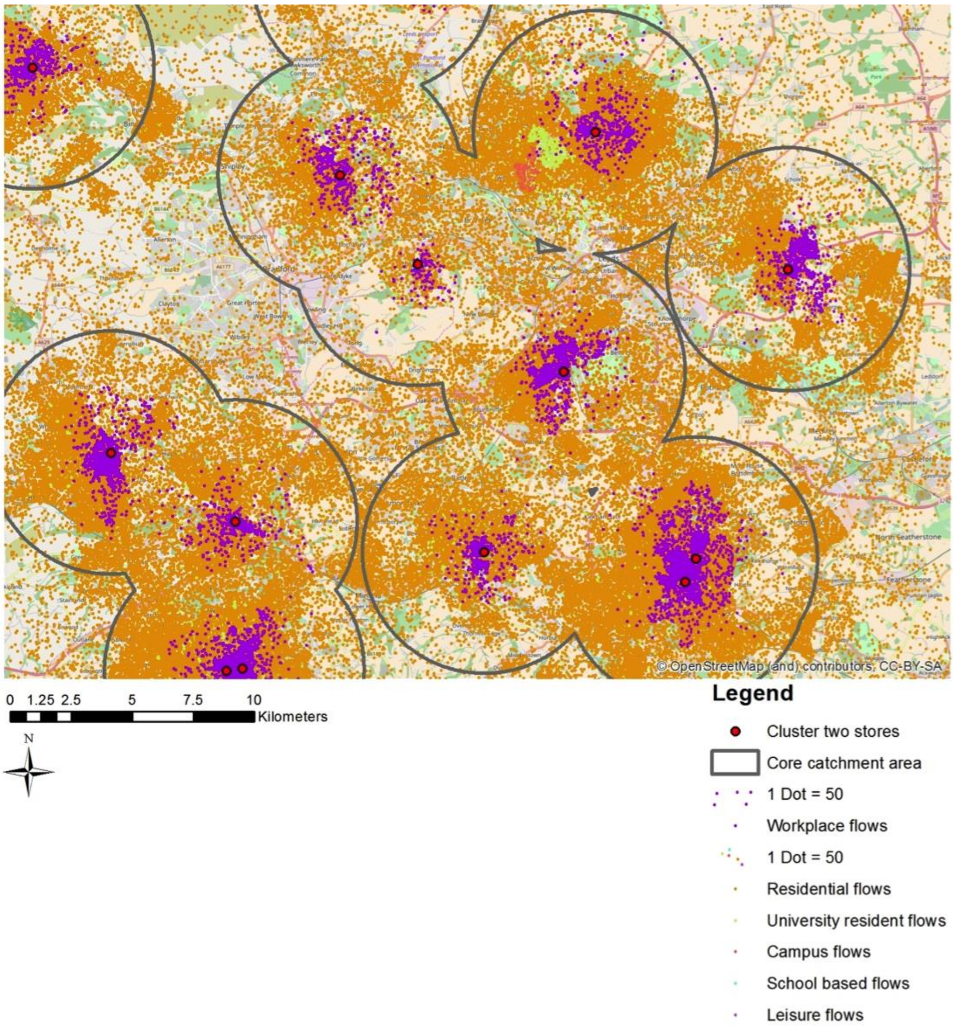
Based on the catchment analysis and widely accepted theories of consumer behaviour, it is expected that distinct shopping patterns will be observed. Again for demand types associated with a lower mobility and low propensity to travel, e.g. school, university or workplace based demand, it is expected that flows will be limited to stores within a close proximity to their location. As a consequence, flows for these particular demand groups will be localised and concentrated around the nearest store. However, due to the considerable residential population within the core catchment (Figure 8.18), and coupled with the temporal sales profiles (Figure 8.17), it is likely that flows to stores in cluster two will be dominated by residential based demand, which may explain the more constant revenue observed throughout the day. The consistent shopping times over the last two decades (East et al., 1994, Ipsos-RSL, 2003) and larger basket size at these stores (Figure 4.8 and in Section 4.5.1) indicate that behaviour in these stores has remained relatively similar, continuing to follow a weekly 'one-stop' shop (Akehurst, 1984, De Kervenoael et al., 2006, Hunter, 2004). As noted in chapter two, there are many components of behaviour that interrelate and result in varying consumer behaviour (Avery et al., 2013, Solomon et al., 2013, Thompson, 2013). The expected residential behavioural norms regarding an average consumer were indeed seen in the loyalty card data i.e. consumers undertaking large weekly shops, tending to travel to the most attractive destination possible with the lowest travel costs incurred. In regards to residential population behaviour, this suggests a propensity to travel further than, for example, students or workers potentially bypassing nearer, less attractive sites in favour of better, more attractive destinations. Therefore, one would expect to see residential flows from further away. In many cases this may occur even when a local store is available. Needless to say, it is widely noted that residentially based consumers will also use smaller local stores. There is plenty of literature noting the drive for convenience and shopping missions centred around 'top-up' type shops (Birkin et al., 2017, Hallsworth et al., 2010, Hood et al., 2015, Wood et al., 2006). This is accounted for by the predicted residential flows to convenience stores in cluster one and cluster three.

Figure 8.18 - Diurnal population distribution within cluster two core catchment areas, West Yorkshire (dots represent 50 individuals)



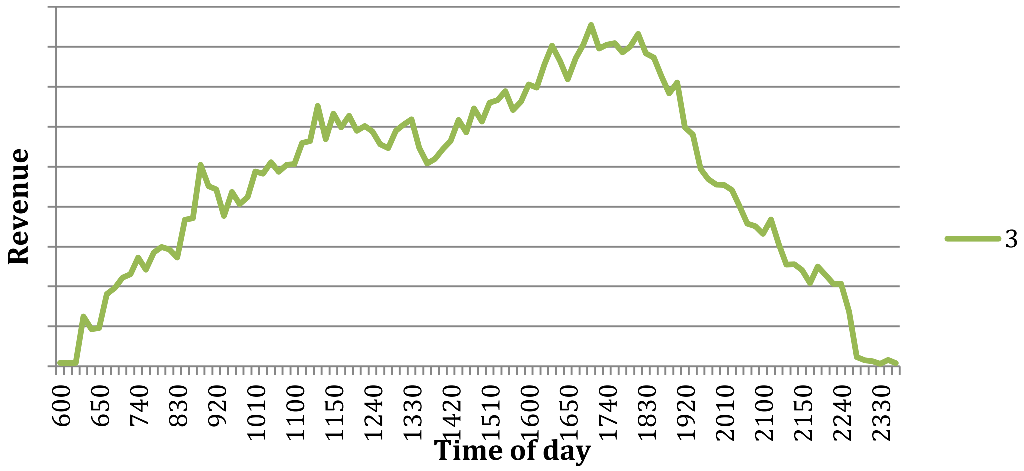
Analysis reveals that the biggest driver of sales in the model for this cluster is residential demand, followed by workplace flows and then university residents. Approximately 80% of the flows to these stores come from within the core catchment area. Those additional sales from outside of the core catchment are mainly residential (Figure 8.19). In these instances stores that are closer are typically much smaller, therefore this behaviour is not unexpected and the model conforms to associated norms expected in supermarket shopping. Overall the behaviour of residential demand when considering their interaction with supermarket stores appears to be consistent. As noted in cluster one, it is clear that model performance is weaker when modelling workplace demand behaviour, allowing distances travelled to stores well above the associated norms, particularly in urbanised areas, densely populated with food shopping opportunities. Needless to say, replicating traditional grocery shopping behaviour from the home appears to be one of the strengths of the model and this is indicated by the better average store sales accuracy predictions for this cluster (Figure 8.12).

Figure 8.19 - Flows (£50 per dot) to cluster two stores by demand group, West Yorkshire



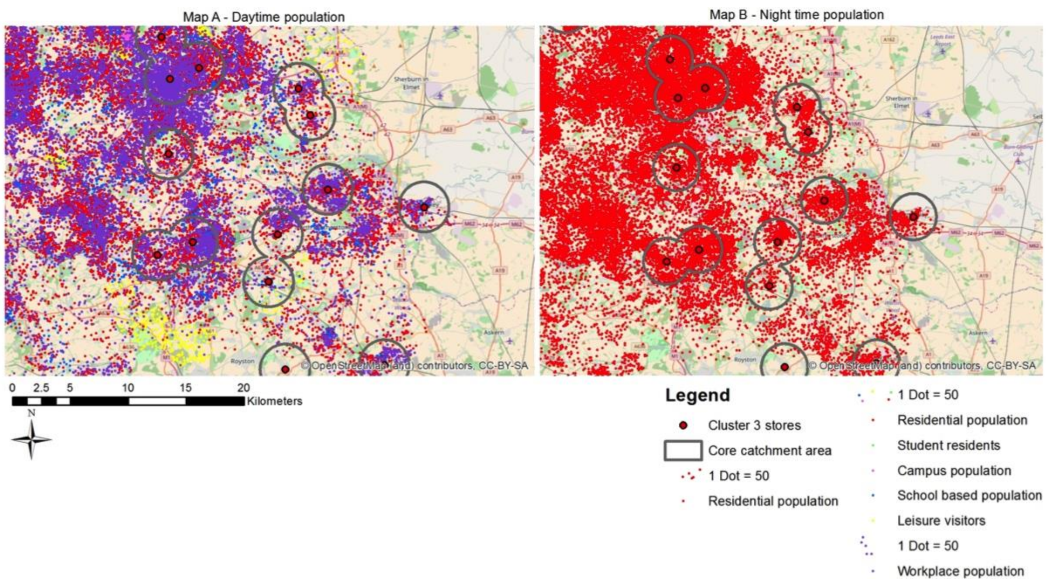
Cluster three stores 'local convenience':

Figure 8.20 - Temporal sales profile for cluster three stores, averaged over a week



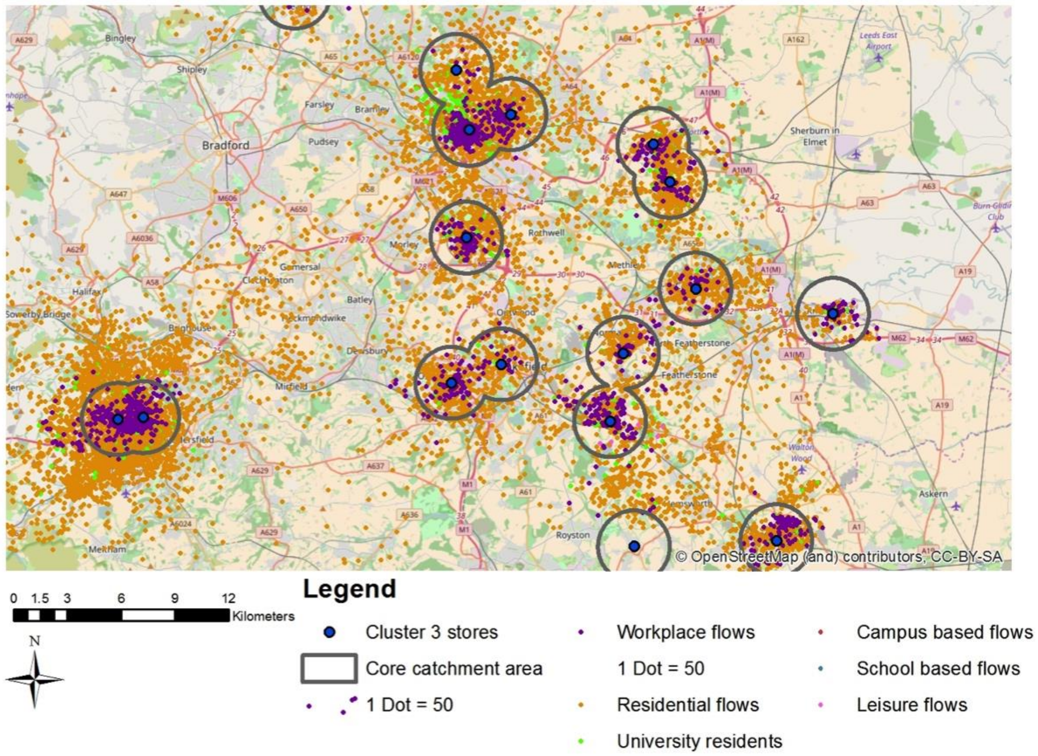
The stores in this cluster are all convenience stores with an average floorspace of 1,808 sqft. They are typically located in secondary or smaller retail centres such as traditional high-streets, suburban or small town centre locations and appear to supply a well balanced demand population in terms of numbers during the night and day. Within the core catchments of cluster three stores, the demand present at certain times of day, while remaining relatively stable, does fluctuate, incorporating residential demand at all hours, as well as students and a workplace population during the day. While there is a net increase within the core catchment population during the day (this is particularly illustrated by a few stores in Leeds which experience high numbers of temporal workplace populations), several core catchments for individual stores appear to demonstrate an outflow of consumers. This reflects the nature of cluster three stores, which represent a ‘local convenience’ type store catering to a mix of daytime populations, but also night time residential populations in predominantly suburban neighbourhoods. It is in these areas where the apparent outflow of consumers appears strongest. This is likely to be due to many of the residents being constrained by daily spatiotemporal activities, potentially resulting in this spatiotemporal pattern (Figure 8.21). This apparent spatiotemporal pattern of the core catchment population is reflected in the observed sales profile. The uplift in sales during the morning and at lunchtime may be the results of customers leaving for work, or purchasing on the way, as well as purchasing their lunch. Similarly, the evening uplift may be the result of sales made on the way home or made locally once at home at the end of the day. Unlike in cluster one, which was predominantly a workplace driven cluster located in city centres, there are fewer periods of prolonged decline in sales. This is to be likely the result of the more consistent presence in residential demand at all times, which was less prevalent in the workplace dominant catchments of cluster one.

Figure 8.21 - Diurnal population distribution within cluster three core catchment areas, Leeds and Wakefield (dots represent 50 individuals)



Based on the apparent diurnal makeup of cluster three core catchments (high levels of residential demand during the night accompanied by an inflow of workers during the day), the observed temporal profile is not unexpected (figure 8.20). As indicated, there are (although lower in magnitude then cluster one stores) morning and lunchtime peaks and there is also an evening peak, which in this cluster is the most profitable period of the day. These times coincide with the behaviour of the temporal profile of the core catchment population, i.e. indicating the presence of temporally fluctuating workplace demand, for instance. East et al. (1994) suggest shoppers will tend to fit shopping around other temporal factors, such as work as theorised. There is, similar to cluster two, a continuous stream of revenue which may be associated with a residual residential population during the day. This particular demand group (potentially retirees or the unemployed, for instance) has been observed to be less confined by temporal constraints (Kenhove and De Wulf, 2000) and so may explain the presence of sales occurring outside of peak periods, e.g. lunch time on a typical working day. Sales at these stores also occur later in the evening, outside of the typical 9-5 working day. This evidence strengthens the assumed classification of cluster three stores, that they operate as ‘local convenience’, and mainly service the surrounding residential population. These indicators of consumer behaviour, which are based on the times sales are made, the spatiotemporal population and the types and locations of the stores, appear to demonstrate that the behaviour within the cluster three stores is similar to the expected behaviour typically associated with the demand type present. Therefore, work based demand should occur locally and should almost entirely occur well within the core catchment area. Berry et al. (2016) suggest that this is typically within a radius of 500m. However, some of cluster three stores are more ‘rural’ than others in the study region and so grocery options may be more limited for workers, so in these instances greater travel distances may also be observed (Wood and Browne, 2007). Residential demand will be responsible for a sizable proportion of the store sales and this should occur relatively locally as well and will be the result of the day and night time population. However, unlike in supermarket stores, the distance travelled by consumers to these stores should be smaller (Wood and Browne, 2007). Nevertheless, these stores, as indicated in chapter four, appear to act as a local convenience store servicing surrounding residents and therefore should attract localised residential expenditure. However, due to their size (a cluster average of 1,808 sqft), consumers are likely to shop only at these stores when more conveniently located. Furthermore, Wood and Browne (2007) note that for convenience store types there will be less emphasis on car borne trade and therefore the spatial extent of consumer interaction will typically be more localised around each store.

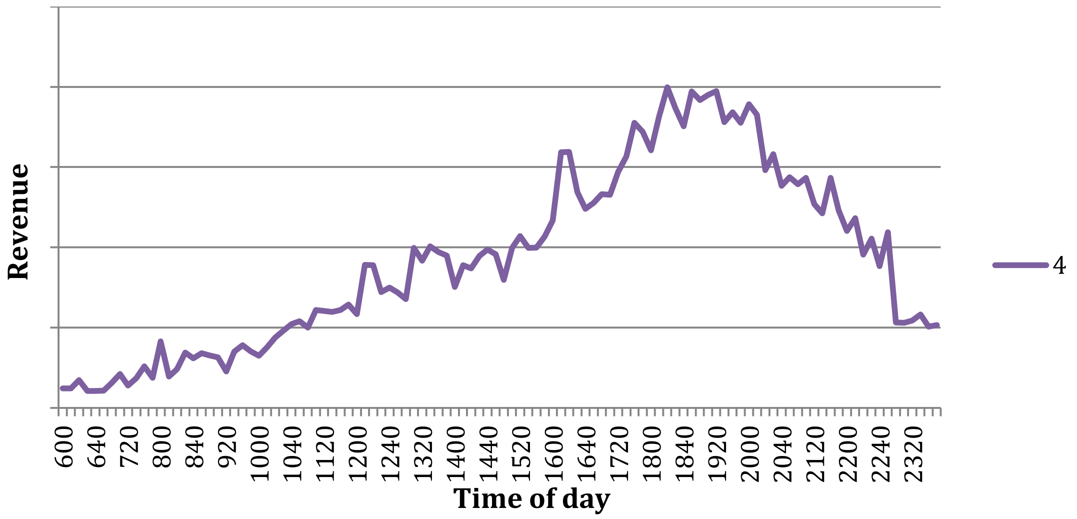
Figure 8.22 - Flows (£50 per dot) to cluster three stores by demand group, Leeds and Wakefield



Following an evaluation of the spatial pattern for flows, it appears that the model typically conforms to the expected normal behaviour of consumers, i.e. residential demand travels furthest, yet remaining relatively local and dominating sales again. However, they do appear to travel greater distances than typically expected. Additionally, workplace based flows were highly localised and also contributed substantially to overall revenues. Correspondingly, the main source of predicted revenue is residentially based, trailed by workplace demand, which was expected, as indicated by the sales profile and spatiotemporal population (figures 8.20 and 8.21). However, the actual model predictions indicate a different story in terms of the model performance. The average store prediction of cluster three stores is approximately 80%, though this is affected by some outliers. There appear to be several stores with considerable under and over- predictions. Some of these major under-predictions can be attributed to missing demand due to their location on the edge of the study area. Those stores located well within the study area that are under-predicting appears to primarily occur when a larger, more attractive store is located in or close to the store core catchment area in question. This was corrected by making adjustments to the propensity to travel of consumers, and could also explain why other stores are over-predicting (when there is less competition present). Over-predictions likewise appear to occur due to an under-sensitive propensity to travel. Subsequently, and as indicated in the previous clusters, whilst the model is capable of replicating differences in spatiotemporal demand type behaviours, predictions were limited by an over generous propensity to travel which was then correspondingly adjusted.

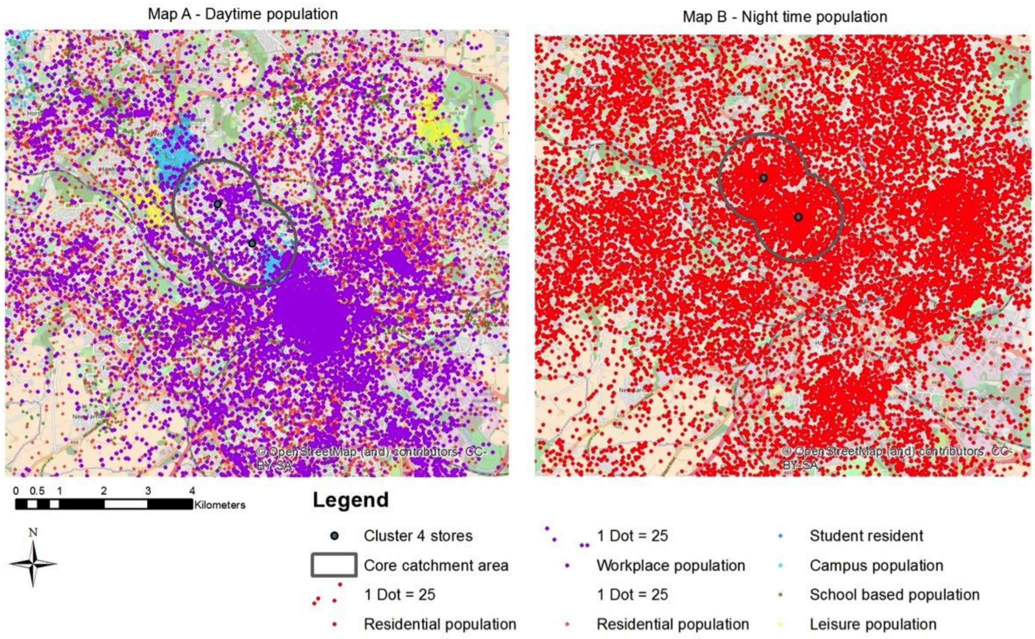
<u>Cluster four stores '*student central'*:</u>

Figure 8.23 - Temporal sales profile for cluster four stores, averaged over a week



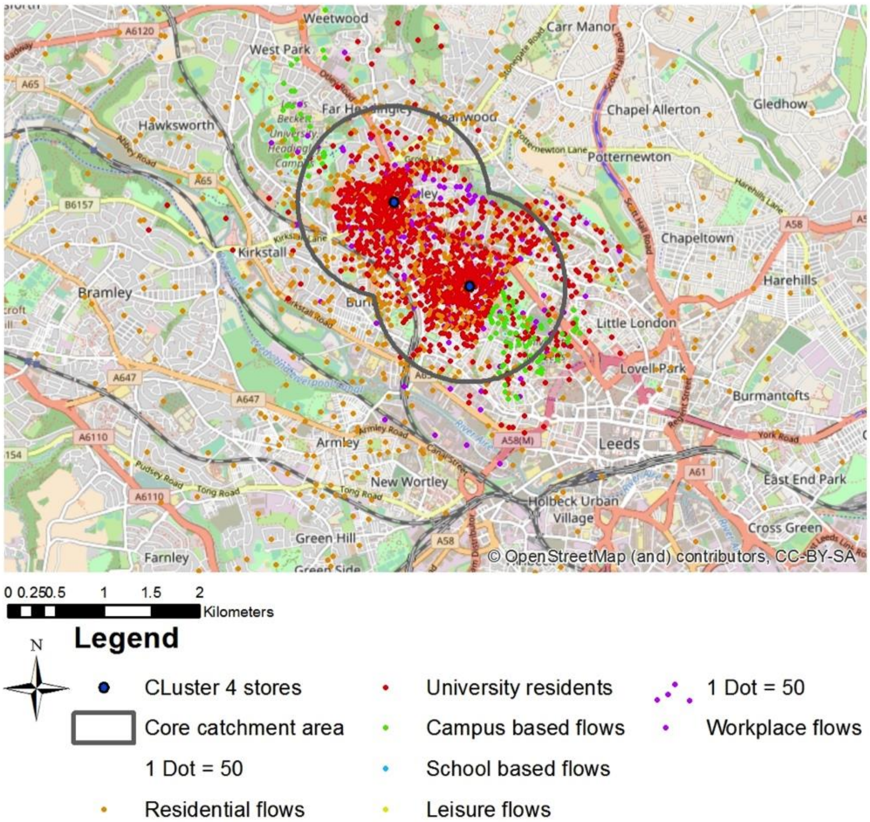
The two stores in this cluster are both located in areas of Leeds dominated by high student numbers. The average sales profile was unique to this cluster and as such they were classified separately from the others very early on in the clustering process. The time profile demonstrates an unusual daily sales pattern with the busiest period of sales during the evening. This is likely to be a direct result of both the spatiotemporal patterns of student populations and their behaviour. The two stores are located quite close to the city centre as well as being between two of the main university campuses of the city. During the night the core catchment area of 1km has a high population as indicated in Figure 8.24 (b). Of this population, over 60% of the residents within 1km of the store are university students. During the day the core catchment area is less populated. Many of the student residents will attend university and while these are relatively close, amassing large numbers of students throughout the day, they are on the fringes of the core catchment and appear to have little impact on store sales. This is based on the apparent minimal sales at the stores during the day when compared with the volume of potential customers at the university campuses for the same time period. There is also an inflow of workers during the day and being near the city centre, there are particularly high concentrations of workers just outside the core catchment area.

Figure 8.24 - Diurnal population distribution within cluster four core catchment areas, Leeds (dots represent 25 individuals)



Within the cluster core catchment area, there are four spatially and temporally fluctuating demand types: residential, student residents, campus based students and workers. Flows decline quite quickly as the distance to each store is increased. Due to the stores being convenience stores, it is unlikely that residents patronise only these stores and so it is likely that only a proportion of the available expenditure in the area will go to these stores. Student residents will make up a sizeable proportion of the revenue generated from within the core catchment. Although Ness et al. (2002) observed that students will sometimes travel further when access to a car is available, many students indicate a tendency to shop closer to home. Therefore, the spatial pattern should be quite well defined and closely situated around each store. Typically, student behaviour is associated with grocery shopping close to home (Devine et al., 2006). Ness et al. (2002) note that student populations as grocery shoppers have received little attention in academia: and even less is known about their grocery shopping behaviour while on campus. While Charles-Edwards & Bell, (2013) have undertaken studies accounting for their spatiotemporal activity during the day, and Devine et al. (2006) report student opinions about campus based opportunities, noting the belief that food was expensive and a lack of healthy food provision on campuses, there are few studies explicitly discussing grocery shopping behaviours on campus. Pelliter & Laska, (2013) looked in detail at dietary and consumption habits of students on campus, but again little is known about the spatial patterns and campus shoppers' behaviour. Here, the expected behaviour is partially derived from their typical shopping behaviour from the home and the temporal sales profile analysis (Figure 8.23). As noted, there are considerable campus based populations during the day on the boundary of the core catchment areas. Furthermore, it is likely that a large majority of the campus-based population will be on foot. It is likely that the propensity to travel will be diminished even further, with students unwilling to travel too far from campus during the day. This assumption was supported by the lack of daytime sales in the cluster four stores, suggesting that campus based students are unwilling to travel more than 500m. Consequently, there should be few transactions generated by campus-based students over this distance, with those that do occur potentially the result of purchases made while travelling. Workplace flows should follow the expected norms previously discussed: flows should likewise originate locally, typically within 500 meters of the store. The core catchment areas of cluster four, when compared to surrounding localities, has a lower workplace population and therefore the volume of flows should be proportionally small as well as concentrated within the core catchment area.

Figure 8.25 - Flows (£50 per dot) to cluster four stores by demand group, Leeds



The model appears to currently replicate the expected behaviour of students, who chose to shop at these stores, whether grocery shopping occurs from their place of residence or from campus in terms of the distance travelled, improving the realism of the these store revenue estimations. As shown in figure 8.25, these demand group flows are confined predominantly within the core catchment. It appears therefore that in this cluster their behaviour conforms to the expected norms, although in previous clusters students were observed to travel further than expected, indicating a need to improve their behaviour within the model. In the previous clusters, it was observed that residential and workplace flows are likewise potentially travelling too far. As the core catchment area for cluster four is the smallest of all the clusters, it is therefore unsurprising to see flows from outside this area.

In summary, the model appears to distinguish the behaviours of different spatiotemporal demand types well. With the data available the expenditure estimates and the expected location of demand types at certain times of day, within the scope of this research, model simulations are far more representative of actual spatiotemporal fluctuations of demand than the pure night time model. The model handles the distinctions between each demand type well, although the cluster-by-cluster analysis (above) identified some weaknesses in the model's behavioural parameters, the evidence points to an improved capacity due to the addition of spatiotemporal catchment areas. Generally, although store predictions are not perfect, the inclusion of spatiotemporal components has improved store predictions overall and subsequently the potential operational and strategic benefits to retailers emphasize the need to not only understand spatiotemporal components, but also indicate the importance of on-going research with implications for commercial applications. In many cases stores that were under-predicting have increased revenues and are now closer to observed store revenue. Similarly, the magnitude of over-predicting stores in most cases has reduced, resulting from the more realistic distribution of daytime expenditure throughout the region. Following further behavioural adjustments to the spatiotemporal demand, such as propensity to travel and brand preference for different temporal demand types, as identified above, it is likely that the improvements to store level predictions accounting for spatiotemporal components when compared to the night time model would become even more apparent. Importantly, the ability of the model to account for the differences and the effects on store sales resulting from different spatially and temporally constrained demand clearly indicates the importance of time for retailers. Hence, understanding how catchment areas fluctuate over various temporal scales will be commercially interesting with potential benefits resulting from the generated insight, thus making this research and investigation of spatiotemporal demand and the related impacts on store sales timely and indispensable.

8.5 Conclusions

The analysis reported in this chapter sought to provide a novel contribution to the disaggregation of different demand types and their behaviours via spatiotemporal components. This has been demonstrated in section 8.2 through the identification of unique behavioural traits, patterns of grocery shopping and expenditure estimates for novel demand types such as university students, school pupils and visitors using a combination of literature and observed consumer data for different times of day. This chapter then sought to demonstrate that it is not only possible to integrate temporally constrained demand layers into a SIM, but that through the incorporation of temporal components of demand the SIM accuracy for store level revenue predictions could be improved, which this chapter has demonstrated. The temporal SIM equation, defined in section 8.3, illustrated that SIMs are capable of temporal refinement, adding in and supplementing different demand types dynamically and by artificially representing different temporal scenarios by separating the model into individual components this was possible. Subsequently, the impact to stores by spatial and temporal drivers of trade can be replicated within the model environment allowing the modelled interaction between supply and demand to have a far more realistic representation of observed market conditions (as demonstrated in section 8.4).

Nevertheless, while it has been possible to demonstrate an increase to the overall predictive capacity of the SIM built in this research via temporal refinement, the model accuracy at a store level remains limited, many of the revenue estimations fall outside of the industry standard of +/-10% of reality. Predictive performance was better and more consistent for supermarket stores than for convenience stores, although the greatest degree of accuracy improvements to store estimations were typically observed in convenience stores. This is likely to be due to the new model which incorporates novel demand type behaviour and better represents the origins of consumers, especially those on foot and in close proximity to stores at certain times of the day. Traditional residential demand alone would not have been able to replicate these conditions, which convenience stores sales are known to be highly dependant on (Wood and Browne, 2007). It is possible that the limited accuracy in store level predictions in the temporal SIM are in part due to its susceptibility to the widely acknowledged pitfalls of SIMs, particularly when trying to estimate convenience store revenue within a SIM environment. The limitations of SIMs and considerations for future research are discussed further in chapter ten. Due to the high usage of SIMs by retail location planning teams (Reynolds and Wood, 2010, Wood and Reynolds, 2011a) the SIM was chosen as the platform to demonstrate the importance of temporal components and the potential improvements to the representation of demand for location planners. In this regard, the objective has successfully been achieved. The insight into novel demand types and their relationship with temporal components of revenue (illustrated in chapters four and five) highlights the necessity for retailers to take into account the temporal components of demand, which have been proven to impact store sales. The next chapter takes the spatiotemporal components and applies the model in a practical setting to demonstrate the commercial application of spatiotemporal data and the increased insight into store revenue estimations.

Chapter 9 - Spatiotemporal ‘What-if’ modelling

9.1 Introduction

This chapter presents the application of the spatiotemporal data in a commercial context for location planning teams and demonstrates the increased detail and novel applications for store revenue estimation techniques utilising spatiotemporal data. This is achieved through a series of spatiotemporal scenarios, which demonstrate the impacts on existing store revenues resulting from spatiotemporal demand. This is then supplemented with several new ‘what-if’ store simulations, estimating potential new store revenues using the custom built spatiotemporal SIM to illustrate the impact of different spatiotemporal demand, and the beneficial insight gained from using spatiotemporal demand layers (which commercial teams can exploit). The chapter concludes with the development of a novel store revenue estimation technique, which using spatiotemporal data (via the cluster groups defined in chapter five) is capable of predicting the typical temporal sales profile of a store over an average day of sales. This offers novel and actionable insight with operational implications for store management, such as adjustments to staffing levels, restocking and the types of products on offer and the promotions on offer at certain times of day, which are traditionally made following a period of review after a new store has opened. This new development offers location planners predictive insight into the typical peaks in sales that can occur at a store allowing for strategic decisions to be made in advance and the optimisation of store operations prior to opening.

9.2 Spatiotemporal ‘what-if’ scenarios

Previously in the research, variations in sales and consumer demand throughout the day have been demonstrated, and implicitly a relationship between temporal demand and store revenue appears to be evident. The effects appear to influence the times that stores generate revenue, suggesting that store performance varies at certain points during the day and is linked to the temporal makeup of demand within the core catchment area. Therefore, it is useful for store location planners to account for temporal fluctuations as the temporal variation in demand and thus, store sales, can considerably alter indicators of store performance. The effect of time is demonstrated below, using the stores in the model that are currently predicting within an acceptable accuracy threshold. First, by illustrating store revenue and demand at certain points throughout the day, it is possible to demonstrate the times that different demand types are typically freely available to undertake grocery shopping in a diurnal sales scenario: in other words when they are not constrained by temporal controls, such as work or school. This demonstrates when various sources of expenditure become important and illustrates the extent to which the sample stores generate potential sales by each spatiotemporal demand type at each point in the day. An illustration of the temporal pattern of the grocery market is likely to be of great interest to location planners and store managers, indicating temporal insight that may be of potential strategic and operational importance. Following this, a long term temporal scenario around the impact of removing or making different demand expenditure available for stores sales will be simulated, replicating various real-world temporal periods, such as school and university holidays. Subsequently, by using and understanding the temporal components of demand, this research demonstrates the potential opportunities from exhibiting novel temporal insight through spatiotemporal scenario modelling of demand over various time scales, i.e. daily and monthly scales, thus not only demonstrating the relationship between store sales and time, but also the potential for commercial applications when presented with detailed temporal data.

To begin, this period of store operation was then segmented into a series of shorter periods throughout the day. Again, these were logically chosen, so that they fit with traditional daytime activities to ensure a fair representation of the day (i.e. morning, lunchtime, afternoon and evening) and demand availability, was subsequently distributed across these. Certain demand types, for instance school and workplace, are typically constrained by environmental factors such as average workday hours or school day times. Hence, in this instance, shopping activity was framed around their typical temporal activity patterns; thus, morning, lunch and afternoon. Daytime leisure demand is somewhat constrained by attraction opening times. Therefore, it is logical that shopping activity will occur more throughout the middle of the day, allowing time for travel to and from the destination. Residential demand is available throughout the day, accounting for those that stay at home, are on holiday or unemployed and retired and are thus freely available to engage in grocery shopping at any time. Residential based university demand appeared to demonstrate a tendency to primarily undertake their grocery shopping in the evening (see figure 8.23). Campus based behaviour within this research is associated with lunchtime purchases. Charles-Edwards and Bell (2013) suggest that this time of the day appears to represent the busiest period of the day, when the highest proportion of students are on campus and thus are not at home at these times. It also appears to represent the time of day least likely to be restricted by formal teaching, as indicated previously by Tomlinson et al. (1973), in addition to lunch time campus based food purchases being shown to be a common trait of university populations (Pelletier and Laska, 2013). Furthermore, the campus based expenditure has been designed to represent daytime purchases of food to supplement, not replace, grocery shopping at home and thus, shopping activity is constrained within the lunch time period.

As noted, the stores selected for spatiotemporal scenario modelling were picked to, meet an accuracy threshold criteria, i.e. store predictions needed to fall within a 10%+/- accuracy range of observed store revenue. This resulted in a total of ten stores, with clusters one, two and three all being represented. Cluster four, which only contains two stores and is dominated by a local student population, did not fall within the selection criteria and is therefore not represented in the following scenarios. Nevertheless, there is a good representation of the other clusters within the 10 sample stores (Table 9.1) and there is evidence of all the temporal demand types within the selected store clusters, as indicated below in Table 9.3, providing a fair representation.

Tables 9.2 and 9.3 below demonstrate the estimated availability of each demand type throughout the day for each store. This is indicated via the 'average temporal constraints' section, whereby demand has been indicated to be potentially available to engage in grocery shopping at times marked with an 'X'. Table 9.2 demonstrates the predicted store revenue as a percentage of total store revenue by stores generated by each demand type. As indicated, revenue has been equally distributed across the temporal availability for each demand type to illustrate the potential volume of sales that may occur at various times during the day. The impact of these fluctuations varies for each demand type, but it appears evident that temporal activity (i.e. specific periods of shopping activity) contributes to total store revenue. For example, store 4725 appears to generate approximately 56% of total store revenue from workplace and university and campus based demand at varying points throughout the day. This insight is important to consider as, for example, store revenue will vary considerably at different times of the day, days of the week and times in the year because of those demand types, and will thus impact the management and operation of this store. For instance, during university holidays and national holidays, the potential shortfall in store revenue for this store could be £19,700, resulting from the missing temporal sales generated from university students and workers. The impact of this type of temporal scenario is demonstrated further below. Table 9.3 adopts a similar framework, except in this instance it illustrates the total grocery expenditure of each demand type within the core catchment area of each store. Again, the total expenditure has been equally distributed to further demonstrate the potential temporal nature of consumer spending, suggesting that various portions of grocery spending are available at certain times of day and that this also varies by demand type. The expenditure values (£) do not represent the sales generated by each store; instead the decision to use total grocery expenditure was decided, as it indicated the sizable volumes of money that grocery stores are able to access throughout the day and their variable nature. For example, in the core catchments of store 2004 and 4218, approximately £179,000 and £94,000 respectively become available from campus-based populations at lunchtime alone. This highlights the importance of accounting for temporal factors in store operation in order to cater for and capture temporal fluctuating demand.

Table 9.1 - Cluster group allocations by store

793 – 2	2004 – 2	4719 – 3	Cluster 1 – Workday convenience
814 – 2	4218 – 1	4725 – 1	Cluster 2 – Traditional supermarket
867 – 2	4480 – 1	4754 – 3	Cluster 3 – Local convenience
4780 – 3			

Table 9.2 - Percentage of total revenue by store, generated by each demand type by time of the day over a week*.

Store ID	Demand	0600-0900	0900-1200	1200-1330	1330-1500	1500-1700	1700-2000	2000-2400
Average temporal constraints	Residential	X	X	X	X	X	X	X
	University					X	X	X
	Campus			X				
	Leisure		X	X	X			
	School	X		X		X		
	Workplace	X		X		X		
793	Residential	12.7	12.7	12.7	12.7	12.7	12.7	12.7
	University	0	0	0	0	1.89	1.89	1.89
	Campus	0	0	2.09	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	1.75	0	1.75	0	1.75	0	0
814	Residential	14.5	14.5	14.5	14.5	14.5	14.5	14.5
	University	0	0	0	0	0.67	0.67	0.67
	Campus	0	0	0.01	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	1.89	0	1.89	0	1.89	0	0
867	Residential	14.7	14.7	14.7	14.7	14.7	14.7	14.7
	University	0	0	0	0	0.47	0.47	0.47
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	1.75	0	1.75	0	1.75	0	0
2004	Residential	12.5	12.5	12.5	12.5	12.5	12.5	12.5
	University	0	0	0	0	1.46	1.46	1.46
	Campus	0	0	2.16	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0.05	0	0.05	0	0.05	0	0
	Workplace	0.67	0	0.67	0	0.67	0	0
4218	Residential	8.97	8.97	8.97	8.97	8.97	8.97	8.97
	University	0	0	0	0	2.65	2.65	2.65
	Campus	0	0	0.15	0	0	0	0
	Leisure	0	0.44	0.44	0.44	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	7.86	0	7.86	0	7.86	0	0

Store ID	Demand	0600-0900	0900-1200	1200-1330	1330-1500	1500-1700	1700-2000	2000-2400
4480	Residential	12.9	12.9	12.9	12.9	12.9	12.9	12.9
	University	0	0	0	0	1.3	1.3	1.3
	Campus	0	0	0.1	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	4.1	0	4.1	0	4.1	0	0
4719	Residential	12.2	12.2	12.2	12.2	12.2	12.2	12.2
	University	0	0	0	0	0.7	0.68	0.678
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	4.6	0.0	4.6	0.0	4.6	0.0	0.0
4725	Residential	5.12	5.12	5.12	5.12	5.12	5.12	5.12
	University	0	0	0	0	4.7	4.7	4.7
	Campus	0	0	15.8	0	0	0	0
	Leisure	0	0.63	0.63	0.63	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	8.59	0	8.59	0	8.6	0	0
4754	Residential	12.4	12.4	12.4	12.4	12.4	12.4	12.4
	University	0	0	0	0	2	2	2
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	5	0	5	0	5	0	0
4780	Residential	10.5	10.5	10.5	10.5	10.5	10.5	10.5
	University	0	0	0	0	0.7	0.68	0.68
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	5.08	0	5.08	0	5.1	0	0

*According to the typical times of day that demand expenditure is available (without constraint) for grocery shopping. The total revenue (as a percentage) has been distributed equally according to the temporal availability for each demand type.

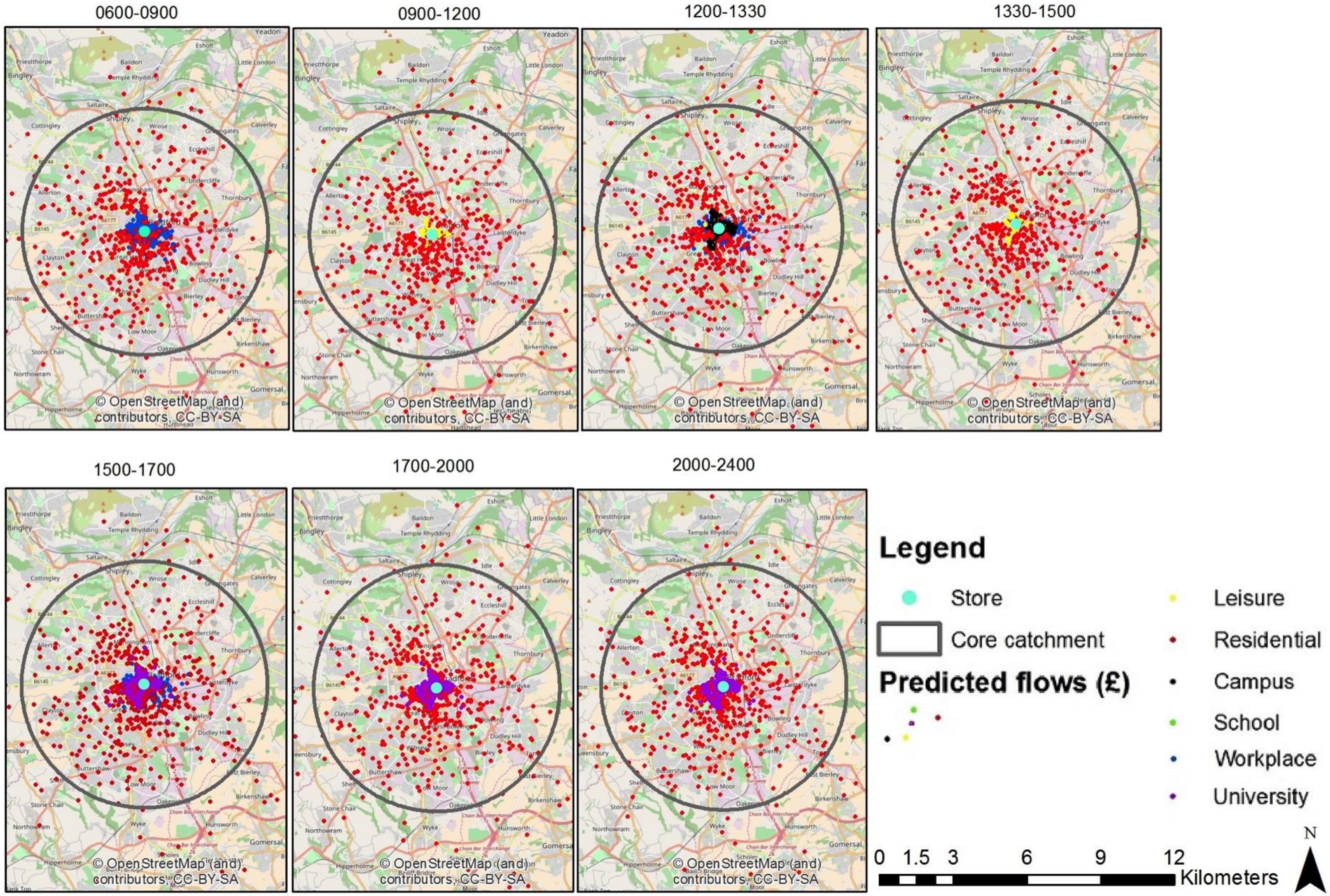
Table 9.3 - Total available expenditure (£) by demand type and time of the day within store core catchment areas for grocery shopping*.

Store ID	Demand	0600-0900	0900-1200	1200-1330	1330-1500	1500-1700	1700-2000	2000-2400
Average temporal constraints	Residential	X	X	X	X	X	X	X
	University					X	X	X
	Campus			X				
	Leisure		X	X	X			
	School	X		X		X		
	Workplace	X		X		X		
793	Residential	543068	543068	543068	543068	543068	543068	543068
	University	0	0	0	0	81273	81273	81273
	Campus	0	0	57388	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	3405	0	3405	0	3405	0	0
	Workplace	115867	0	115867	0	115867	0	0
814	Residential	681332	681332	681332	681332	681332	681332	681332
	University	0	0	0	0	57888	57888	57888
	Campus	0	0	10250	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	6120.3	0	6120.3	0	6120.3	0	0
	Workplace	158965	0	158965	0	158965	0	0
867	Residential	389258	389258	389258	389258	389258	389258	389258
	University	0	0	0	0	20228	20228	20228
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	2779	0	2779	0	2779	0	0
	Workplace	65752	0	65752	0	65752	0	0
2004	Residential	691394	691394	691394	691394	691394	691394	691394
	University	0	0	0	0	303193	303193	303193
	Campus	0	0	179306	0	0	0	0
	Leisure	0	5656	5656	5656	0	0	0
	School	5345	0	5345	0	5345	0	0
	Workplace	173755	0	173755	0	173755	0	0
4218	Residential	917465	917465	917465	917465	917465	917465	917465
	University	0	0	0	0	331875	331875	331875
	Campus	0	0	94185	0	0	0	0
	Leisure	0	6527	6527	6527	0	0	0
	School	3981.7	0	3981.7	0	3981.7	0	0
	Workplace	396718	0	396718	0	396718	0	0

Store ID	Demand	0600-0900	0900-1200	1200-1330	1330-1500	1500-1700	1700-2000	2000-2400
4480	Residential	624164	624164	624164	624164	624164	624164	624164
	University	0	0	0	0	148336	148336	148336
	Campus	0	0	3422	0	0	0	0
	Leisure	0	5341	5341	5341	0	0	0
	School	4690	0	4690	0	4690	0	0
	Workplace	118610	0	118610	0	118610	0	0
4719	Residential	47593	47593	47593	47593	47593	47593	47593
	University	0	0	0	0	2715	2715	2715
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	0	0	0	0	0	0	0
	Workplace	10612	0	10612	0	10612	0	0
4725	Residential	859206	859206	859206	859206	859206	859206	859206
	University	0	0	0	0	124827	124827	124827
	Campus	0	0	33696	0	0	0	0
	Leisure	0	4423	4423	4423	0	0	0
	School	10070	0	10070	0	10070	0	0
	Workplace	231490	0	231490	0	231490	0	0
4754	Residential	87771	87771	87771	87771	87771	87771	87771
	University	0	0	0	0	4010	4010	4010
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	52.333	0	52.333	0	52.333	0	0
	Workplace	8616.7	0	8616.7	0	8616.7	0	0
4780	Residential	89842	89842	89842	89842	89842	89842	89842
	University	0	0	0	0	3421.3	3421.3	3421.3
	Campus	0	0	0	0	0	0	0
	Leisure	0	0	0	0	0	0	0
	School	18.333	0	18.333	0	18.333	0	0
	Workplace	15803	0	15803	0	15803	0	0

* Total available expenditure has been equally distributed according to the typical temporal availability of each demand type (without constraint). Note: the expenditure values represent the total available expenditure on food and drink by demand type within each store core catchment and not the revenue actual store revenue.

Figure 9.1 - Spatiotemporal pattern of predicted flows (£) by demand type and by time of day to store 4725



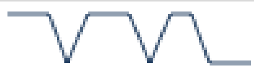
The implications of time on store revenue are demonstrated further below through potential real world temporal scenarios. In the following scenarios, the impact of demand availability (and as such flows to a store and thus store revenue) is demonstrated. The outputs generated demonstrate the effects of store level revenue experienced at these times, highlighting the importance of understanding temporal demand types for existing store revenue and also when considering new developments. Figure 9.1 represents a spatial visualisation of a typical daily temporal scenario (shown in Table 9.2), demonstrating the spatiotemporal pattern of temporal flows (£) to a single store. In this example, it is possible to observe the spatial pattern of sales generated by each demand type by time of day. It is clear that the accumulation of sales and the source of store revenue fluctuate throughout the day. The incorporation of spatiotemporal dynamics has clear advantages over traditional location planning models utilising only residential populations. Not only does it increase the capacity of the spatial representation of demand, locating consumers in the right location at the right time, but the disaggregation of demand caters for the more complex behaviours exhibited at certain times of day and by certain demand types, that a residential population alone cannot account for. In this instance the clear and highly localised and temporal constrained flows generated by certain demand types, e.g. work, leisure and university demand at certain points in the day, appear to play an important part in the subsequent store revenues, which again would be likely to be missed when using a residential night time population model only. Incidentally, the spatial pattern and the temporal distribution of flows adopted in Table 9.2 for this store appear to correlate with the parent cluster mean sales profile (figure 8.14). The data informing the sales profile was taken from observed temporal sales data. The apparent similarities in terms of store sales and demand availability, i.e. between the observed mean sales profile and the predicted temporal flows depicted in Figure 9.1, thus appear logical and offer confidence in the temporal framework and resulting impacts on store revenue. Based on the diurnal scenarios, it is evident that sales fluctuate throughout the day and how the source of revenue varies in terms of demand type. Using the scenarios, it could be possible for retailers to investigate the impact of changes in revenue experienced by adapting temporal scenarios, e.g. the proportional drop in sales on a weekend when it is likely the work and school based demand will be missing or the impact of one off events, such as football games, adding extra demand into the leisure layer. The evidence in tables 9.2 and 9.3 and Figure 9.1 highlights the obvious importance of the temporal activity of demand over short-term time scales, e.g. a diurnal scale. The obvious benefits from representing demand in a store's core catchment area temporally is that revenue predictions are not only likely to be more realistic, but also that models can be adapted dynamically to account for real world temporal situations with the insight potentially aiding strategic and operational decisions for both the existing store network and potential extensions.

The long term implications for store revenues experienced by spatiotemporal demand have previously been suggested: for example, the impact of seasonal tourism was identified by Newing et al. (2013a) and appeared to illustrate clear impacts on store revenues at certain points throughout the year. Similarly, in cities and regions with high volumes of students at university (which varies at certain times of the year), the potential grocery shopping market can be sizable, representing millions of pounds in expenditure (Ness et al., 2002). The region of West Yorkshire is a prime example of this, with Leeds LAD containing three universities and several thousand students. In many instances students are only in the place of study during term time, normally returning home during the holiday periods. This can have a serious impact upon expenditure levels with a potential decline in store sales experienced at a local level. Furthermore, the longevity of these lower sales periods can be extensive. For example, the University of Leeds had 21 weeks of non term-time in 2016 (University of Leeds, 2017). Potentially, in stores that cater to a large proportion of students, which has been shown to occur in the cluster four stores, these periods of likely reduced demand result in large periods of low revenue, resulting in stores under-performing for long periods and considerable decline in store performance. Below are the outputs of store revenue predictions simulating the long-term temporal impact of university term times (Table 9.4): weekly store revenue is simulated during term-time and during holidays to demonstrate the immediate impact on store sales on a weekly scale. This is then scaled to illustrate the potential decline in sales resulting from holiday periods throughout the year. This is supplemented with the potential decrease in store performance (trade intensity - TI), in addition to an accumulated total of sales that stores could potentially lose out on during the holiday periods. Finally, the magnitude of the impact of temporal student demand throughout the year on store revenue is illustrated using the Excel Sparkline feature and is illustrated at a store-by-store level.

Immediately obvious from the scenario is that although sometimes the sales driven by students are low, due to the long-term temporal components of student demand, e.g. resulting from long holidays throughout the year, all stores experience effects on store revenue to some degree. This is shown via the Sparkline's, which fluctuate throughout the year with varying degrees of magnitude. Needless to say, store 4725 experiences the biggest temporal effect from the sample of stores, with all things being equal, from student demand which represents a total of 30% of the predicted store sales for that store and experiences a potential fall in TI of over £10 per sqft during the holiday periods. For this store, this represents a significant change in store performance with the store potentially underperforming for a least 1/3 of the year. However, the total decline in potential sales are highest in stores 793 and 2004. In these instances, the sales driven by student demand only represent 8% and 6% of total store sales respectively and so proportionally the effects are lower. However, the annual impact in terms of the potential volume in sales lost during the holiday periods is much higher; (£1.4 and £1.1

million each) and thus may represent a far more important commercial issue for the retailer as a business. The three highlighted stores from the temporal scenario illustrate the importance of understanding temporal aspects of catchment area dynamics. Clearly, the model predictions suggest that spatially the temporal effects can vary considerably, impacting stores to varying degrees and this is why it is important to disaggregate demand temporally and to consider the spatial distribution too. Furthermore, over long periods the accumulated fluctuations in sales can result in large drops in store sales. These insights will likely be of interest to retailers and will be reported back to the retail partner. Understanding that a store might underperform for parts of the year could influence the operational management of the store during those times, e.g. reducing staffing levels or the delivery of stock. Likewise, the volume of sales change will also be of interest. While from a store level perspective, a decline in sales is always important, the relative impact across the national store network may be minimal, e.g. store 4725 and likewise, vice versa, e.g. 793, which will likely be of interest to head office. These insights become even more important to retailers if they can predict them in advance, see section 9.4.

Table 9.4 - Temporal scenario model predicting the impact of long-term temporal fluctuations in student demand on store level sales

	Term time - weekly average			Holiday time - weekly average		Accumulated impact of holidays	Annual impact from holiday periods
Store ID	Student driven sales (%)	Student driven sales (£)	Total Store revenue	Change in TI (£)	Total Store revenue	Total potential drop in sales (£)	Academic year starting September
793	8	71,067	891,193	-1.12	820,126	1,492,407	
814	2	7,816	334,964	-0.26	327,148	164,136	
867	1	8,670	567,976	-0.23	559,306	182,070	
2004	6	56,410	744,193	-1.30	687,783	1,184,610	
4218	8	5,015	46,648	-2.80	41,633	105,315	
4480	4	2,577	59,754	-1.01	57,177	54,117	
4719	2	576	23,850	-0.37	23,274	12,096	
4725	30	14,714	34,095	-10.35	19,381	308,994	
4754	6	2,145	32,197	-1.62	30,052	45,045	
4780	2	772	28,818	-0.59	28,046	16,212	

9.2.1 Summary

This section has presented both long and short-term temporal scenarios for fluctuating demand, proposing the resultant potential impact on store level revenue occurring at certain temporal situations. The predictive outputs appear to demonstrate evidence of considerable sales fluctuations relating to the temporal availability of demand. The analysis demonstrates that store core catchment areas are highly dynamic, experiencing varying demand both spatially and temporally, and appear to correlate with observed data supporting the findings. Furthermore, temporal changes to demand appear to have potentially major implications for stores, which could have implications for store operations and new store development. Thus, the analyses demonstrate the importance of considering location planning as a dynamic process whereby available expenditure can fluctuate. Models that fail to account for temporal components will fail to capture a true representation of demand and consequently, store predictions and consumer flows and our knowledge on temporal activity are likely to be a poorer representation of observed conditions. The scenarios reveal that the magnitude of the effects on store revenue will vary by demand type, by time of day and by the store. Therefore, the evidence indicates that it is important to consider future temporal situations, not only because trade will vary over time, but because the implications for retailers could result in significant financial consequences at the store level. The promising and logical outputs from the temporal scenarios appear to suggest that temporal scenario modelling will be beneficial to retailers for forecasting stores: it can provide future insight into store performance via detailed simulated temporal activity of different demand types. The monitoring of existing stores is one component that could benefit from temporal analysis, but it is likely that new store development would benefit, providing predictions of the financial viability, store performance and also revenue stability of a new store over various temporal scales. This level of insight could provide retailers with warnings of future peaks and troughs in sales at certain times of the day or the year, allowing them to prepare for, and improving comprehension of, variable store trading patterns thus helping to understand why and when particular patterns are observed.

9.3 New store revenue modelling

The aim of this section is to now apply the spatiotemporal SIM in a typical commercial situation (new store development) and perform the role of store revenue estimation which would support store location planning. The aim is to estimate the revenues of four ‘new’ stores for our retail partner and to demonstrate the increased detail and insight that can be achieved by using spatiotemporal demand layers. This provides novel analysis regarding store revenue predictions and it allows retailers to better understand the source of consumer demand, the volume of sales by different demand types and to infer the times of day sales are likely to occur (predicting the time of sales at a store level is considered further in the next section). The advantages for retailers from the additional level of information regarding store demand (provided from the spatiotemporal SIM) offer a range of positive benefits; location planners can make better informed decision making via more realistic demand layers due to the spatiotemporal distribution of consumers and thus, a better understanding of core catchment area characteristics and better predictions as more elements of demand are considered. Practical implications resulting from a better understanding of the characteristics of the local demand (i.e. are they workers, residents students or visitors); can affect the store layout and the types of products and promotions on offer i.e. lunchtime meal-deals, newspaper in the mornings, or alcohol which as noted by Berry et al. (2016), experienced an increase in sales in some workplace zones at the end of the working day. Furthermore, it also offers insight into how demand will vary throughout the day based on the type of demand present in the store's catchment areas and their individual temporal activities (as illustrated in section 9.2). As a result this strategic insight can improve the decisions made by retailers regarding a store's operation, for example, stores can adapt their product displays or adjust staffing levels to ensure they cater to the needs of the consumers throughout the day (Waddington et al., 2017). Three of the 'new' stores used for 'what-if' analysis (convenience stores) were based on actual store openings undertaken by the retail partner since the data for this thesis was provided. The final store (a supermarket) was located in an area with a limited presence by the partner and other competing retailers. The location of these stores is shown below in figure 9.2 and is accompanied by a summary of the store details (Table 9.5).

Figure 9.2 - Locations of 'new' stores

Table 9.5 - Summary of 'new' stores

Location	Format	LSOA	Store size
Vicar Lane	C-Store	E01033010	1591
Woodlesford	Supermarket	E01011637	10000
Bingley	C-Store	E01010578	2303
Royal Park Road	C-Store	E01011445	2628

The store revenue estimations and flow maps for each of the 'new' stores is demonstrated below. The outputs are broken down by each demand type to demonstrate the varying levels of sales driven by each group of consumers. This is compared with the revenue estimations of a SIM using only a residential (night time) population, which subsequently highlights the increased level of detail regarding demand expenditure offered by a model with spatiotemporal layering. Table 9.6 shows the volume of estimated sales (£) to each of the proposed stores for the residential only model and then for each demand type in the spatiotemporal demand model. Figure 9.3 demonstrates the proportion of total estimated sales by each demand type in each store. Discussed further below, figure 9.3 in particular highlights the impact of temporal demand types on store revenue and hence, the need to appropriately disaggregate demand so that this can be represented better in store revenue estimation techniques. Figures 9.4 and 9.5 illustrate the spatial pattern of flows (£50 per dot) to each store for both the residential only model (figure 9.4) and then the spatiotemporal model (figure 9.5), where flows are illustrated for each demand layer.

Table 9.6 - Total store revenue estimates (£) for residential only model (left) **and for spatiotemporal model (by demand type)** (right) **for 'new' stores**

Store Location	Residential only	Daytime Residential	School	Visitors	Student	Campus	Work	Total
Royal Park Road	74360	39027	0	0	56168	4066	6575	105836
Vicar Lane	31959	25316	0	370	4205	64	9565	39520
Woodlesford	155784	139739	0	0	2466	0	2260	144465
Bingley	36988	33184	330	0	1223	0	10230	44967

Figure 9.3 - Proportion of the total estimated store revenue by demand type

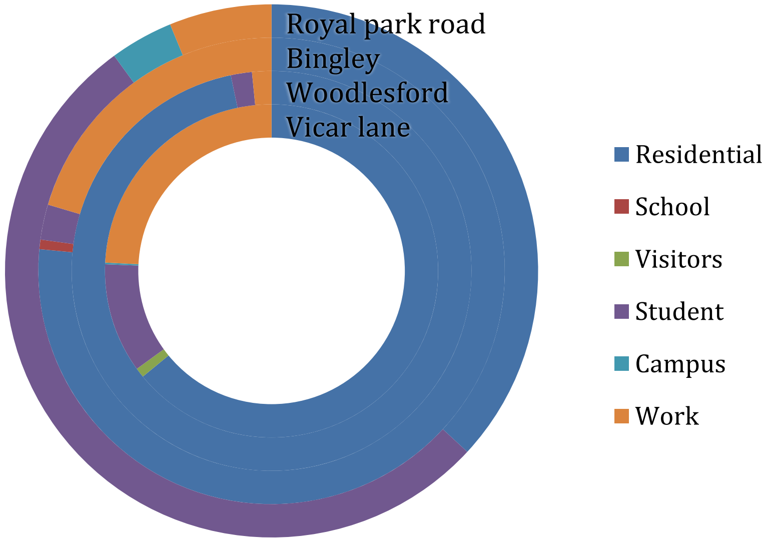


Figure 9.4 - Flow (£50 per dot) to stores for residential demand only model

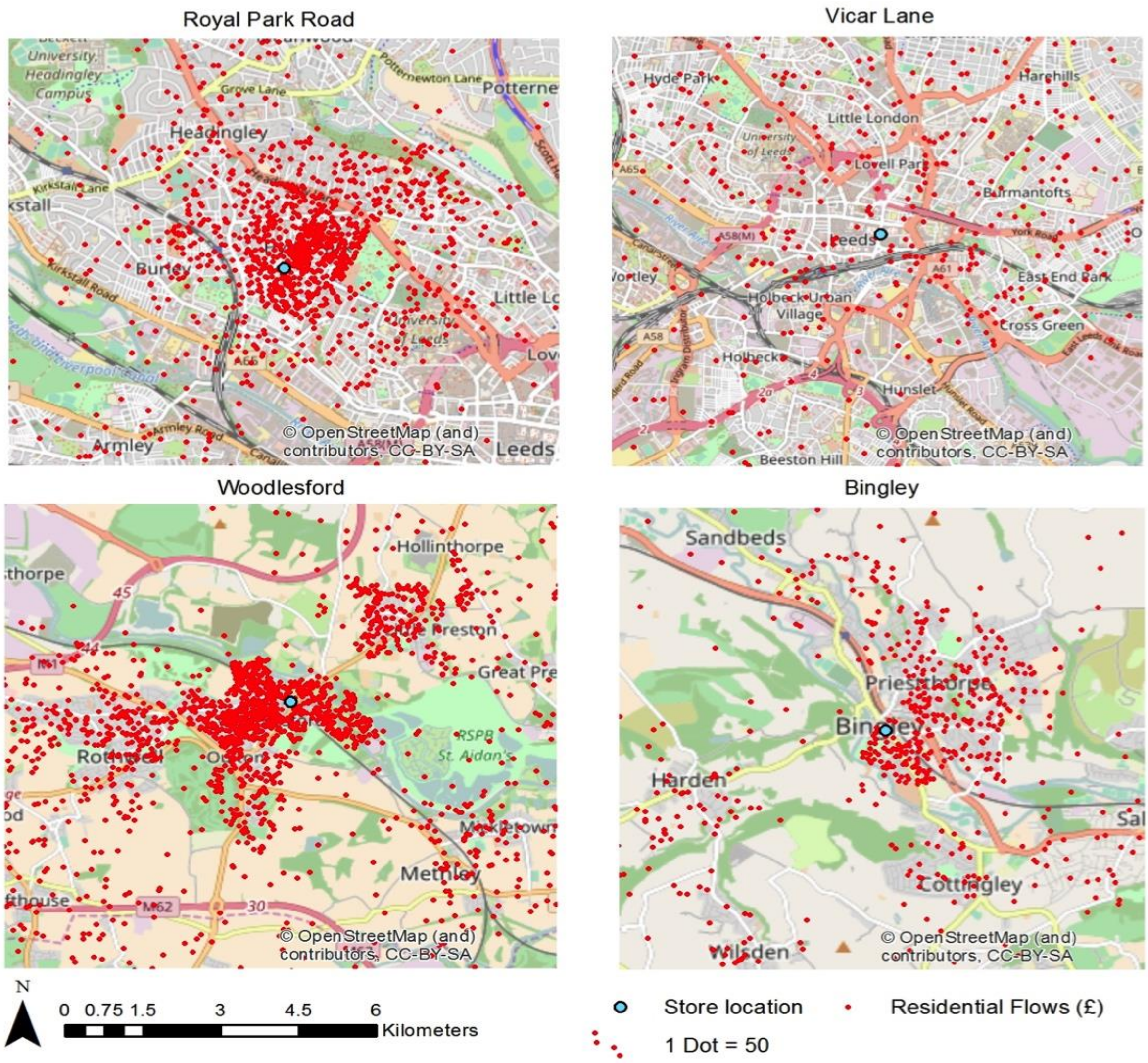
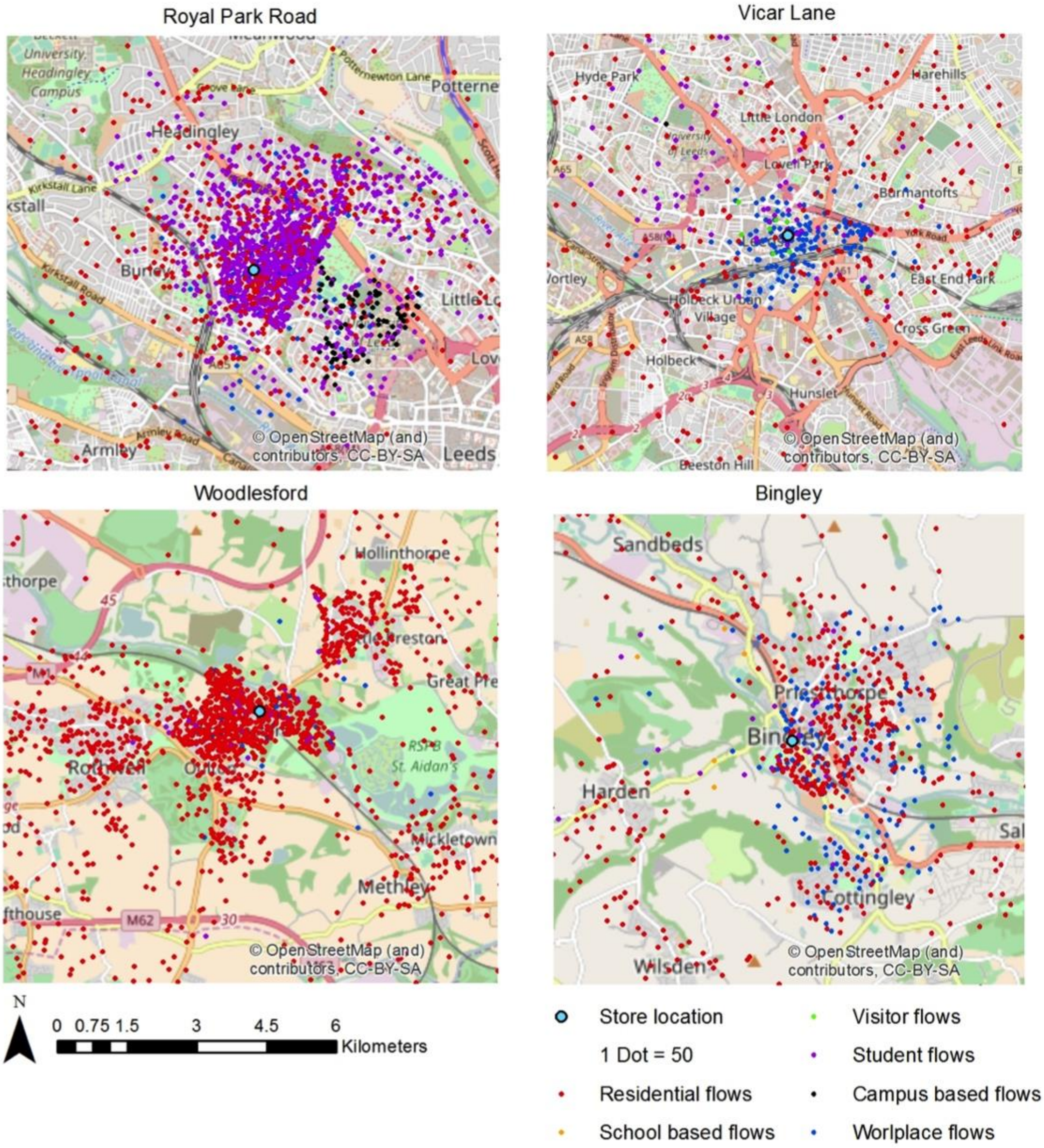


Figure 9.5 - Flow (£50 per dot) to stores for spatiotemporal demand model



The outcomes above, for the what-if 'new' store revenue estimation illustrate the clear differences between, and the power of, a spatiotemporal SIM compared with a residential only SIM. Chapters four and five have previously indicated that census based residential only populations have limited representation of actual daytime consumer populations. This has been indicated in the literature (Bell, 2015, Martin et al., 2013, Martin et al., 2015, Newing et al., 2013a), suggesting that there was the need for an increased level of temporal disaggregation regarding consumer demand and as a result models are far more realistic as a result of their inclusion. The analysis above further strengthens this argument, demonstrating the increased level of detail on offer to location planners when using spatiotemporal demand layers. Figure 9.3 demonstrates the complexities of store revenue highlighting the varied proportions of sales generated by different demand types. The observed differences in each store also further highlight the importance of modelling with spatiotemporal demand because the impact of specific demand types on total store revenue appears to vary spatially and is not consistent across the entire store network. This is illustrated in figure 9.5 where the volumes of flows to each store are clearly influenced by different demand types, which vary in each location. This detail is not present in the residential only model flows (figure 9.4), which in some cases it can be argued results in inaccurate revenue estimations due to the increased flows generated by temporal demand types e.g. visitors and workers at the Vicar Lane store or students at the Royal Park Road store and workers at the Bingley store, which a residential only model cannot simulate.

The analysis highlights that a spatiotemporal model will simulate the flows of expenditure for a daytime population far more realistically than a residential only model. Aside from the improved demand distribution resulting from using spatiotemporal demand layers, from a practical point of view the insight generated by the disaggregation of store revenue by demand type (figure 9.3) can have beneficial implications for store operations. For instance understanding that a certain proportion of revenue will likely be generated by a particular demand type (whose behaviour and temporal activity have previously been discussed), such as students at the Royal Park Road store or workers at Vicar Lane and the Bingley stores and can influence staffing rotas or the types of products on offer, for example. These are all necessary considerations that store managers and head office management will consider at every store, once open. The novelty of SIM developed in this thesis offers, is that it provides not only improved revenue predictions but also insight into these issues in the location planning stage, allowing retailers to prepare and make better informed decision about store operations regarding spatiotemporal demand and sales fluctuations that has previously not been possible with more traditional SIMs as part of the location planning stage.

9.4 Temporal sales profile clusters as a predictive tool

This section continues with the theme of temporal analysis and links to the novel cluster segmentation presented earlier. This section now aims to evaluate whether it is possible to use the original cluster groups and their mean temporal sales profiles, as a crude indicator in attempting to forecast the potential temporal sales profile of other unknown stores. In other words, similar to the analogue model approach used by the grocery industry, the aim is to compare individual stores with characteristics similar to one of the four originally identified cluster profiles (discussed below) and to test whether it is then possible to predict the typical trade profile to expect in a store over the course of a day, i.e. the times that stores generate sales, but not the revenue volume. The predicted cluster's allocation based on subjective analogies will then be compared to a K-means cluster analysis, where stores will be sorted into the original cluster classifications following the original decision tree, using each store's observed temporal sales data. Subsequently, an assessment of the predictive accuracy will be measured, i.e. how often the cluster prediction was correct and thus a similar sales profile was observed, demonstrating the potential capacity of the approach. Any evidence indicating the potential feasibility for using the temporal sales profiles as a predictive tool capable of forecasting the expected temporal store trade patterns could have commercial benefits, leading to an increased awareness of store level trade with implications for commercial operation. If successful, this approach, with further refinement, could offer promising temporal insight for retailers, informing the decision making process influencing store layouts and store operations via the improved level of temporal insight generated from understanding the typical times of day a store potentially generates revenue.

Analogue models are widely used in location planning and in particular by the grocery industry for predicting potential store sales (Benoit and Clarke, 1997, Birkin et al., 2002, Reynolds and Wood, 2010). In brief, the analogue model works by drawing on existing data, looking at a series of key characteristics that classify a store. This store, which typically offers detailed historical revenue data, is then used to predict store revenue for other similar stores. By comparing a new or existing store with the same or similar characteristics; it is then assumed that the existing store revenue is indicative of the potential sales due to the strong similarities (Birkin et al., 2002, Wood and Browne, 2007, Wood and Reynolds, 2011a). Hence, analogies are drawn between a store and another store already within a retailer's network, suggesting that a similar result will be expected. Birkin et al. (2002) and Reynolds and Wood (2010) have previously discussed the analogue technique for location planning in more detail. Reynolds and Wood (2010) in particular provide a more recent background on the current usage and applications of the method by location planners. Widely cited in the location planning literature, and its application in the commercial sector similarly recognised, adopting the analogue model approach helps to improve the robustness of the planned methodology as a commercial toolset. Despite the limitations potentially resulting from subjective decision making processes, allocating a predicted store trade profile to new stores, the numerous previous applications and continued reliance on analogue models for store revenue predictions by industry teams indicate a reliability of the method by the grocery sector and its corresponding predictions.

The classification of stores into the appropriate predicted cluster, and as such the corresponding expected average sales profile, is based on the assessment of criteria for each store. Criteria separating stores within each of the original clusters from on another were identified, drawing on a mixture of physical and trade area characteristics (presented below). In this case, 'new' stores (a previously unused dataset where the stores were not included in the original clustering process, see below) will then be subject to evaluation, attempting to draw analogies with a particular cluster group's criteria. The predicted parent cluster mean sales profile is then assumed to indicate the potential temporal trade pattern expected by the store in question. Subsequently, each of the remaining stores was then grouped into the relevant cluster with the most likely sales profile based on similar analogies. The capacity of this technique as a predictive tool with regards to commercial potential will be assessed following the analogue comparison, using observed temporal sales data for each of the tested stores. Below are the criteria, set within this research, which stores belonging to each of the original cluster groups identified in chapter four typically encompass. Test stores, and thus likely trade profile predication, will be based on analogies constrained within these criteria. The outcome and comparison of the cluster group predictions using categorical data will be indicated below alongside the clustering outputs using actual temporal sales data (Table 9.9).

Similar to the method used in chapter four, test stores were clustered using a K-means cluster analysis using temporal cumulative revenue data. This was presented in 10-minute intervals starting from 6am (the point from which stores begin to open). The point in time at which cumulative revenue, in each store, exceeded each percentage decile, e.g. the time at which revenue exceeded 10%, 20% etc. for each store, was used as the input variables for the cluster analysis. Rerunning the k-means cluster analysis with all the stores available tested the robustness of the original clusters, in regard to the representation of the four clusters and their ability to operate as parent groups for additional stores. It was plausible that with an increased sample size, the original cluster groups may not be representative, resulting in a different grouping of stores through newly identified temporal trade patterns. A k-means cluster analysis was thus rerun using all of the available data. The outcome of this analysis was tested in two stages; the first stage was to review how many of the original 47 stores, using all the store data (a total of 129 stores) fit within the original four cluster groups. The output (shown below in Table 9.7) identified that 42 of the 47 stores, when using the larger temporal dataset, remained in the same classification (89%).

Cluster 1 Stores:

- Highest proportion of daytime populations e.g. workplace population within 500m
- Typically located in town/city centre close to main shopping and business districts

Cluster 2 Stores:

- Typically have the largest amount of available floorspace
- Located at out of town or retail park type location types
- High residential population within driving distance and high secondary workplace population due to surrounding shopping facilities

Cluster 3 Stores:

- A mix of both high and low day and night populations e.g. workplace and residential populations
- Located in smaller towns, residential suburbs or secondary amenities sites

Cluster 4 Stores:

- Very high residential student population within close proximity to store
- Located in student dominated suburbs

Table 9.7 - Comparison of the classification of the old stores, between using only the 47 original (old) stores and all 129 (old and new) stores

47 OLD STORES		Clusters created using all 129 stores – old and new			
		1	2	3	4
Clusters based on 47 old stores only	1		**12**	1	
	2	**14**			
	3		4		**14**
	4			**2**	

Note: Those in **bold** are where stores have remained in the same category – although the cluster number may have changed.

The five stores that did change were as follows; store 4218 (located in Brewery Wharf) changed from the 'workday convenience' cluster to 'student central', which had a higher tendency to demonstrate sales in the evening and may reflect the transition of the catchment resulting from recent student accommodation developments in the area. Likewise, four stores from cluster three, the 'local convenience' cluster, were reclassified into the 'supermarket' cluster according to the times that they generated sales.

Next, the new test stores (a total of 82 stores) were then grouped into the original cluster classifications and this was achieved by applying the decision tree from the original cluster analysis on to the 82 test stores, using the J48 decision tree process in the software package WEKA, which is an algorithm for modelling the classification process of a dataset (Bhargava et al., 2013, Patil and Sherekar, 2013) i.e. it identifies how the cluster groups were split using temporal sales data, and then applies the same rules on the new stores.

The decision tree for the original cluster analysis (47 stores) resulting from the J48 process is as follows:

If 90% reached earlier than 19:00 **class as 2**

If 90% reached after 19:00; if 30% reached before 13:00; if 60% before 17:00 **class as 3**

If 60% after 17:00 **class as 1**

If 30% after 13:00; if 30% before 14:00 **class as 3**

If 30% after 14:00 **class as 4**

This classification of the test stores (using the original cluster groups) was then compared to the clusters derived from the K-means cluster, using all 129 stores as described above. The outcome of this analysis (Table 9.8) found that the original cluster remained the same for a total of 71 of the test stores when classified using all of the temporal data. In other words, the original clusters were accurate 87% of the time.

Table 9.8 - Comparison of the classification of the new stores using the decision tree from the original 47 "old" stores with the classification using all 129 (old and new) stores.

82 NEW STORES		Clusters created using all 129 stores – old and new			
		1	2	3	4
Clusters based on 47 old stores only	1		**15**	2	
	2	**20**			
	3	1	8		**33**
	4			**3**	

Note: Those in **bold** are where stores have remained in the same category – although the cluster number may have changed. 71 of the 82 stores remained in the same classification (87%).

Test stores cluster group allocations via the k-means cluster method were then compared to the predicted cluster group memberships derived from the subjective analogue comparison (which used a mixture of demand and store level data within the constraints of the criteria defined above). The summary of this analysis is presented below in Table 9.9. A full breakdown of the cluster predictions and the cluster classifications using the decision tree at a store-by-store level can be seen in the appendix (Table 12.2).

Table 9.9 - Comparison of test store cluster group predictions using categorical data and observed revenue data.

82 NEW STORES		Predictions following analogical comparison using categorical data			
		1	2	3	4
Clusters based on 47 old stores only	1	**9**		8	
	2	2	**18**		
	3	6	3	**32**	1
	4				**3**

Note: Those in **bold** are where stores have remained in the same category – although the number may have changed. 62 of the 82 stores remained in the same classification (76%).

As a novel predictive tool, analysis of the results indicates that overall this has proved successful, accurately predicting the sales profile of stores (based on the four clusters identified and defined in the research) 76% of the time. Based on the evidence presented here, this suggests that it is not only reasonable to group stores with the four temporal trade profile clusters, but that it appears possible to compare new stores using the criteria also defined in the research to predict the type of temporal trade profile a store is likely to experience 76% of the time. The apparent success as indicated by the data suggests that this approach would likely be adoptable by retailers for use in a predictive capacity and that they could benefit from the novel temporal insight generated at a store level. The biggest source of error was experienced from predicting 8 cluster one stores as cluster three stores and vice versa: 6 cluster three stores as cluster one stores. The source of this error derives from the difficulty in assigning the correct cluster to these stores. This is likely to result from the fact that many of the stores in cluster one and three have very similar store and core catchment area demand characteristics. Subsequently, temporal trade profiles, which have already been demonstrated via the mean cluster profiles (Figure 9.6), are similar, which when combined together, creates issues in the subjective assessment (potential solutions are discussed below). If the rate of correctly predicted stores had been higher for stores in both clusters one and three, then the overall predictive accuracy could have been as high as 95%. That said, as shown in Figure 9.6, due to the apparent similarities in the sales profiles of cluster one and cluster three stores, in addition to the similarities in the original definitions (see section 4.5.1) and assessment criteria above (i.e. morning and lunch time peaks in sales and presence of workplace populations in both clusters), the impact of misidentification may not be of critical importance. Therefore, while generally this has brought the accuracy of the method down, from a commercial interest perspective, the slight differences in predicted store trade profile would potentially have less of an impact on store planning and operational decisions. Furthermore, it is likely through better and more detailed commercial data, such as product purchased by time of day (which retailers will have access to) that this misclassification error could be improved.

Overall, cluster two store predictions were predominately accurate. However, three convenience stores predicted to follow the 'traditional supermarket' trade profile were in fact more like the trade profiles experienced by 'local convenience,' although realistically this is not unexpected, and through refinements to the analogue process (though subjective) this could be improved. Furthermore, the two supermarkets that were predicted to follow a sales profile more similar to the 'workday convenience' trade profile were in fact more similar to the 'traditional supermarkets' cluster. Overall, all the test store supermarkets based on the observed sales data appear to trade similarly to one another, i.e. they are representative of the 'traditional supermarket' average sales profile. This evidence suggests that except in extreme circumstances (as with a supermarket in the original cluster, which was found to demonstrate a 'workday convenience' trade profile) supermarkets appear to almost always follow the same, traditional sales profile, generating revenue at the times shown and have continued to follow a consistent pattern for nearly three decades (East et al., 1994, Ipsos-RSL, 2003). Overall, the cluster three stores ('local convenience') total is the largest cluster group, though this is not unexpected as there has been a dramatic increase in the number of convenience stores in recent years (Hood et al., 2015). Typically, convenience stores are developed to cater to the needs of convenience and offer local grocery shopping opportunities and thus high volumes in this cluster are not unexpected. The stores predicted to exhibit the temporal sales profile of cluster four ('student central') stores, with the exception of one convenience store, were all accurately predicted. The one store that was predicted inaccurately demonstrated a smaller student population, suggesting that the impact on store trade, i.e. becoming dominated by student demand and their grocery shopping habits, is more likely to occur when the core catchment is densely populated by students. Otherwise, the impact created by small student numbers will be minimal and offset by the behaviours and spending of other demand groups. The evidence of a high degree of predictive accuracy for this cluster, in addition to the highly dissimilar, and distinctive nature of cluster four stores' temporal trade profiles, suggests that this approach has a potential to be highly successful for predicting stores that are likely to demonstrate the cluster four sales profile.

For comparative purposes, the observed temporal sales data as a percentage of total revenue (£) by time of day for a sample of successfully predicted stores, one from each cluster group and the mean cluster group profiles, are presented below in Figures 9.6 and 9.7, alongside an analysis of the GOF between the observed sales profile and the predicted typically expected cluster profile (Table 9.10). As indicated, the predicted cluster group and the corresponding typical temporal trade profile for the stores below and their actual observed sales profile have obvious similarities in the times of day that revenue is generated, demonstrating the commercial potential of this method as a predictive tool. While the statistical assessment does not demonstrate a perfect match, this is not unexpected, although the values are relatively high. It would be very unlikely for a test store to have the exact trade profile as the mean cluster to which they belong. Furthermore, this approach was intended to act as a predictive tool to offer an indication of the typical trade profile to be expected. Based on this criteria, the statistical evidence and the observed similarities in temporal trade profiles (Figures 9.6 and 9.7), it appears to fulfil this objective more than satisfactorily. Thus providing a reasonable prediction of the type of trade profile, i.e. the typical times that a store will generate sales throughout a day.

Figure 9.6 - Average temporal trade profiles as a percentage of total revenue (£) for the original four cluster groups

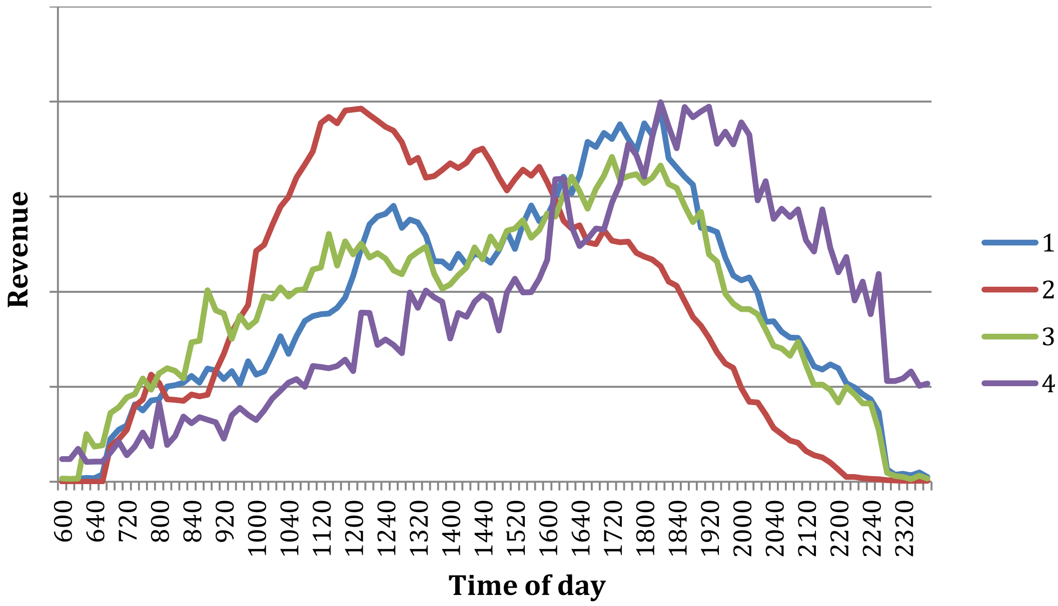
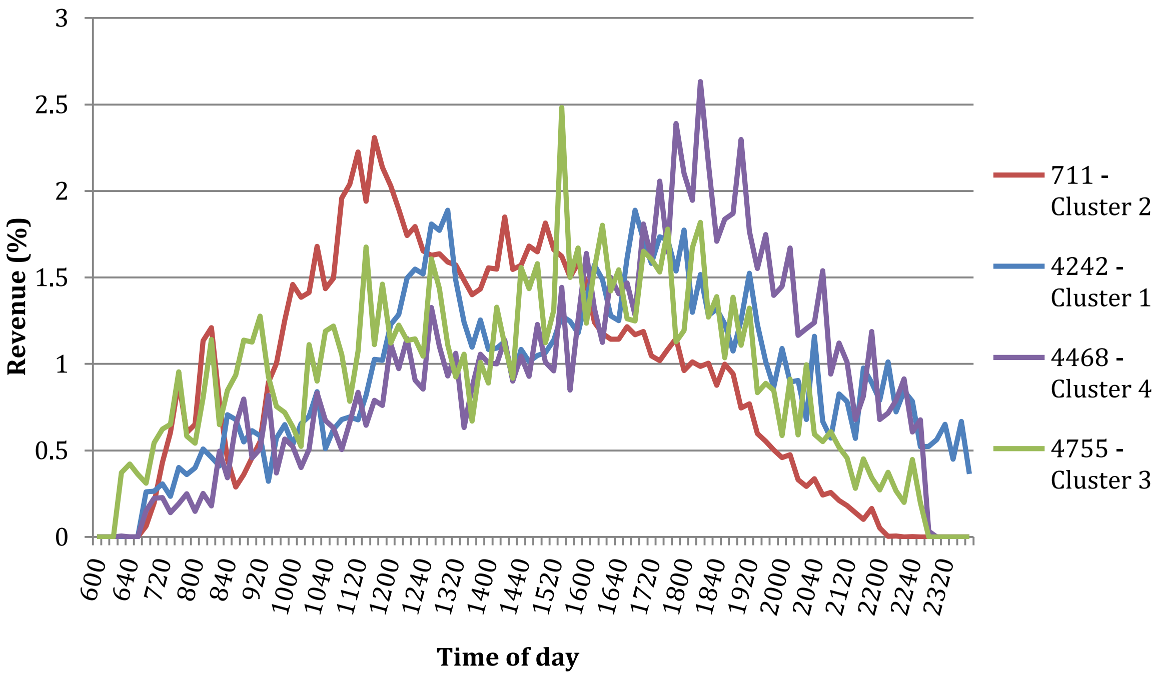


Table 9.10 - Goodness-of-fit assessment between observed sales profiles and mean cluster profiles for sample stores using 104 time points

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Correlation	0.91	0.96	0.88	0.87
RSQ	0.82	0.93	0.78	0.76

Figure 9.7 - Observed temporal trade profiles as a percentage of total store revenue (£) for successfully predicted stores from each cluster group



As mentioned above, there are some limitations to the approach that limited the maximum level of accuracy achieved. Primarily, it appears that these are related to subjectivity and the categorical nature of the cluster criteria and cluster allocation for new stores. The result is that cluster membership boundaries are quite flexible and it can be difficult to categorically state which cluster a store belongs to without a set of quantifiable membership constraints. That said, the criteria were applied logically and supported by a good knowledge of the area and the stores themselves, which appeared to work, though this could be improved. Furthermore, the use of a fuzzy or soft cluster segmentation, such as Gaussian Mixture Methods approach, as opposed to a hard cluster group allocation, such as a k-means clustering method, may resolve this issue when clustering the UK. In this instance, stores which may belong to more than one cluster, which may be the case for cluster one and three, could be assigned proportionally across multiple clusters. Being able to suggest a store may display temporal patterns similar to more than one cluster could help to improve the accuracy of predictions. This clustering method could therefore be applied to future applications, though with the current sample size it was possible to handle this manually. This could allow an analyst to sort a store across more than one cluster, when they appears to be similarities that make it difficult to assign to a single membership, while still providing intelligent insight about potential temporal store behaviour. Likewise, future iterations could additionally benefit from the increased access to commercial data collected by retailers, which will probably offer a far more detailed picture of store and consumer characteristics and behaviour. This is likely to afford a deeper level of insight, in addition to further categorical and potentially quantitative data for stores that could make the process of predicting cluster membership more robust. For instance, the use of temporal sales data alongside the items purchased at certain times of day could reveal interesting insights that could improve the understanding of consumer behaviour at those times of day. Potential patterns could then be a deciding factor in cluster membership. If a store and its catchment area characteristics are very similar, as observed to be the case for many of the stores in cluster one and three, data revealing that the peak in sales at lunch time is primarily ‘food for now’ type products, e.g. meal deals, could suggest that it is more likely to be linked to workplace shoppers than those from home, indicating that this store may demonstrate a temporal sales profile similar to the store in cluster one. It is possible that the increased data would also reveal more intrinsic patterns in the cluster profiles, which could provide quantitative thresholds to improve the assessment criteria. This could allow stores to be sorted according to specific values in the data, removing user subjectivity. However, at present, it is unknown whether this would yield a higher level of accuracy. But it would be wise to trial this approach if sufficient evidence for the four clusters in the data was identifiable. That said, many location planners continue to use gut feeling and experience to deal with store development issues, acknowledging that it is still a vital part of location planning, alongside predictive modelling techniques (Birkin et al., 2002, Reynolds and Wood, 2010), so it may transpire that the analysts’ experience of their customers and store network could be more appropriate. Thus, the results achieved using a subjective approach for predicting cluster membership could be more beneficial and generate more accurate predictions.

Nevertheless, the subsequent benefits to retailers from improving their understanding of temporal components of store and consumer behaviour and predicting trade profiles could provide a competitive advantage, which in the closely fought grocery market of the UK will likely be highly desirable. Particularly for new stores, having a reasonable idea of the times that a store will generate revenue and the accompanying insight about consumer demand, their behaviour and needs would allow retailers to prepare to better fulfil consumer needs. This could be the implementation of operational decisions, translating into staffing and product/restocking optimisation at a store level. This could allow stores to open ‘optimised’ to the catchment area demand, thus running at a higher capacity immediately, instead of adapting over a period of time. Strategic benefits resulting from a better understanding of temporal sales, i.e. understanding when stores perform well and when sales decline, could facilitate better strategic decision making; stores could then take advantage of busy periods, pushing certain products at certain times of day, as well as being prepared to combat periods of lower sales, by offering promotions to drive additional sales, for example. Ultimately, the use of temporal trade profile derived clusters, presented and discussed in this section, has demonstrated a capacity to work successfully as a predictive tool. The strength of the accuracy and supporting statistical analysis demonstrate this potential, not to mention the possible increase in accuracy resulting from future refinements eliminating current weaknesses in the methodology. Therefore, this evidence justifies the methodology, but also demonstrates the potential commercial opportunities that could be gained by location planning teams if they were to adopt this approach. Due to the promising results already demonstrated so far, the findings and the method will be passed to the data partner. To the best of my knowledge, this analysis is novel and attempts to predict the times that stores typically generate revenue has not yet been attempted by the data partner and may prove advantageous.

9.5 Conclusions

The aim of this chapter was to demonstrate store revenue estimation modelling using spatiotemporal data and to forecast the temporal sales profile of a store using the cluster groups developed earlier in chapter four. The first half of this chapter demonstrates the potential novel insight generated from using a SIM, which incorporates spatiotemporal demand layers and the opportunities for novel lines of analyses that this offers to location planners. Section 9.2 demonstrates the ability to simulate the general impact of certain demand types over time, on individual stores revenues, as well as being able to simulate the potential effects in future situations. This was illustrated using a university student demand and demonstrated that stores will experience and uplift and a decline in sales at certain times of the year, impacting store performance. It is likely that this information would not be apparent in revenue estimations when using more traditional demand layers. Section 9.3 demonstrated the application of the spatiotemporal SIM in a traditional capacity, used by location planners for store revenue estimation of new stores. Based on the evidence and previous discussion presented in chapters four and five regarding daytime population distributions, the flow maps for each of the stores (figure 9.5) resulting from the spatiotemporal model are far more representative of the likely observed origin of flows experienced by the stores. Furthermore, the insight into the proportion of sales at a store being driven by individual demand types (figure 9.3) by the SIM could help retailers adapt stores to maximise profits through more targeted store missions. The updated SIM better accounts for the different behaviour and movements of daytime populations, and the differences compared with the residential only flows (figure 9.4) are clearly apparent and appears to represent a marked improvement regarding accurate daytime consumer geographies. As previously noted, three of the proposed stores were based on actual store developments undertaken by the data partner. It is hoped that through the continuing relationship with the retail partner and ongoing research, that the ‘new’ store revenue estimations simulated in this section will be compared against observed store revenue data, to assess the performance of the model in its intended capacity.

The latter half of the chapter demonstrated that the mean sales profiles from the clusters defined in chapter four could be used to estimate the typical sales profile of new stores. The earlier analysis indicated evidence of clear temporal patterns in store sales, likely caused by the relationship between stores and the temporal dynamics of core catchment areas resulting from the types of demand present. This insight suggests that if the membership criteria for each cluster group could be identified it was plausible that a new store's temporal sales profile (using the mean cluster group sales profile) could be predicted by drawing analogies and sorting the store in to one of the four cluster groups. This proved not only possible, but also that it could be achieved with a relatively high degree of accuracy. The original clusters were shown to be representative of the wider store population 89% of the time and the cluster predictions had a success rate of 76%. The temporal pattern of actual observed sales data for a sample of tests stores was then compared with the mean cluster groups and appeared to demonstrate clear similarities and was supported by the GOF analysis in Table 9.10, further suggesting that this objective was successfully achieved. The overall accuracy of the method could be further improved through a greater and more robust set of membership criteria, in particular for cluster one and three stores, which had very similar characteristics. The stores 'incorrectly' sorted into one of these clusters were the main source of inaccuracy. This could be achieved by including temporal product sales data, comparing the types of products bought at certain times of day and could offer greater insight about the types of consumers shopping at stores. This purchasing behaviour may offer sufficient detail to distinguish store membership between a cluster one store, dominated by workplace behaviour and a cluster three store, with a high focus on supporting local neighbourhoods. Had this been implemented at this point in time (this was not possible within the scope of this research because that data was not provided) the predictive accuracy could have been greatly improved. This novel analysis, using temporal data, has produced an output with practical benefits and implications not just for improving academic understanding but also with clear implications for location planning teams as well as store management. In times of economic uncertainty, such as the present, store expansion programs will likely slow down and there will be an increased focus by location planning teams on the drivers of trade surrounding existing stores. Understanding the complex and varied temporal nature of core catchments will likely form an increasing component of commercial analysis, and being able to predict the timings that individual stores are likely to generate revenue will likely be an attractive prospect to retailers. This study has shown that stores and models are impacted by temporal aspects of consumer behaviour and that novel and competitive insight can be generated from temporal data. As noted the insight presented in this chapter can have clear practical implications not just for academia but for retailers too, with benefits that will impact the operational and strategic planning of everyday store operations. The practical impact and potential outcomes for retailers using a spatiotemporal SIM is discussed in the concluding chapter.

Chapter 10 - Conclusions

10.1 Introduction

The main objective of this thesis was to model spatiotemporal fluctuations of consumer demand in the UK grocery sector, and analyse the impact on retail store sales, to use within a well established predictive modelling technique (SIM) currently used by industry. The work reported in the thesis demonstrates the successful completion of this goal; the researcher demonstrates that the findings of this research will have an on-going impact for future academic research in store revenue estimation, but also practical implications for future location planning analyses. This chapter concludes this thesis by summarising the main research findings and contributions of this research. In the next few sections the main aims presented in chapter one are addressed and will be discussed as follows: section 10.2 will cover the analysis of spatiotemporal store sales, section 10.3 will address the analysis and findings relating to spatiotemporal demand and section 10.4 will summarise the findings of store revenue estimation techniques using spatiotemporal components. In addressing each of the aims set in this thesis, the methodologies and data sources are also critiqued. This is followed by a section reflecting closely on the key contributions of the thesis for location planning in industry, discussing the potential benefits and practical implications for retailers and location planning teams. Section 10.7 will suggest possible research agenda for future work, which is followed with some concluding remarks that reflect on the overall achievements of this thesis.

10.2 Summary of spatiotemporal sales

The first key aim of this thesis was to investigate spatiotemporal fluctuations of store sales in the UK grocery sector at a store level. Chapter four introduced the temporal aspects of store level sales and demonstrated sales fluctuations occurring throughout a typical day, but also temporal variance occurring on different days of the week, in different store formats and different location types in response to demand side factors. This was completed by using commercial temporal transaction data collected by a major UK grocery retailer which represented a rare opportunity for academia, generating novel insight (Birkin et al., 2010, Newing et al., 2014b). The impact of temporal demand side factors upon store sales were likewise a focus in chapter five, where for the first time in the literature, observed temporal consumer data (in the form of loyalty card data) was compared alongside observed temporal sales data. Comparisons were made across both chapter four and five contributing to a novel discussion concerning the temporal drivers of temporal store sales variations. To date there have been few attempts to explore the patterns of temporal store sales in detail using data provided by retailers, and to relate patterns with potential demand side drivers. This research represents one of the few studies to attempt this. Newing et al. (2014) and Berry et al. (2016) highlight the importance of understanding temporal fluctuation in store sales, noting that temporal demand types such as visitors and workers can have a considerable impact on stores sales at certain points in the day or year. From a practical point of view it is important for retailers to understand the temporal dynamics of a store so that they can optimise store operations appropriately. Preceding this analysis, little was known about the impact of consumer behaviour on sales at different times of day (Waddington et al., 2017). This research has helped to fill this current gap in the literature, as well as providing practical insight with actionable impacts for commercial location planning teams. The research findings relating to temporal sales data accumulated in a unique and novel cluster segmentation presenting temporal fluctuation occurring at different types of grocery stores. The ongoing focus throughout the entire thesis on the spatiotemporal analysis of store sales opens new lines of investigation for commercial location planners, through a better understanding of the grocery market with opportunities for novel predictive analysis.

The completion of this research aim was possible through the rich data source provided by the data partner, which offered 24hr transaction records for a single week of sales in 136 stores throughout Yorkshire and Humberside. A sub-sample of 48 stores located within West Yorkshire was used for the analysis. The analysis of temporal sales was first presented by store format, comparing the average sales profile throughout the day for supermarkets and convenience stores. Store sales were shown to fluctuate throughout the day, both demonstrating periods of uplifts and reduced sales. The major difference was that supermarket sales primarily occurred throughout the middle of the day and convenience stores peaked in the evening. Prominent studies from the last two decades offer historical insight into supermarket sales throughout the day and on different days of the week (East et al., 1994, Ipsos-RSL, 2003) which this research builds on. Not considering the increase in sales that occur on Sundays (owing to the more liberal Sunday trading laws introduced in the 1990s), supermarket sales have remained relatively consistent over the last 20 years (table 4.3) suggesting that supermarket consumer behaviour has changed very little, with sales predominately occurring midday and with Saturday being the busiest day. Little is known about previous patterns of trade occurring at convenience stores, as data relating to convenience stores is only recently becoming available (Hood et al., 2015, Wood and Browne, 2007). Therefore, the temporal analysis of convenience store sales contributes new insight, building on the retail literature surrounding supermarket sales. The analysis presented in chapter four offers academia possibilities to understand the pattern of sales at these stores and sheds new light on the impact of consumer behaviour. Convenience stores remain relatively consistent throughout the week, though Friday appears dominant. Sales throughout the day peaked in the evening, though smaller peaks were also observed during the morning and at lunch times. The apparent lack of these peaks on the weekend (figure 4.7) suggest that these may be linked to temporal demand only available during the weekday, such as workplace consumers, school pupils or commuters and to 'top-up' type shopping behaviour. Berry et al. (2016) observed similar findings suggesting that stores were likely to exhibit similar peaks in trade at those times in stores located in close proximity to workplace populations. Their research also noted that many of the stores had very low residential populations in close proximity, further highlighting that store sales estimations are likely to remain limited if they do not focus on temporal components, such as transient populations, leading to under-predictions and potentially illuminating why conventional SIMs have traditionally not been used to predict convenience store revenue.

The fluctuation of store sales became more apparent when sales were compared using the partner's in-house location types, illustrating the average sales profile of stores within each location type. As well as noting differences occurring throughout the day, and between different location types, it became evident that there were clear trends observed in some of the store sales with peaks and troughs occurring at certain times of day and in certain location types, suggesting that similar conditions of drivers of trade could be occurring. Blythe (2013), East et al. (1994) and Solomon (2013) note that consumer behaviour typically follows habitual patterns, and that shopping times are often constrained by static external factors. There seemed strong evidence which indicates that observed sales patterns (at certain times of day) are the result of specific demand types that follow, or are constrained, by spatial and temporal circumstances, the result being that store revenue and the time that sales are made are dependant on the makeup of a store's core catchment area. Thus, understanding when stores make money and subsequently the temporal drivers of trade is crucial to our understanding of consumer behaviour, the impact on store revenue, and revenue estimation. The impact of spatiotemporal demand on store sales is likely to continue to be important to retailers due to consumer lifestyles in the UK becoming increasingly irregular i.e. fewer people are constrained by the 9-5 lifestyle (Hallsworth et al., 2010).

A cluster analysis was run to provide further insight from the temporal data and identified that four distinctive cluster groups could be observed in relation to the times the individual stores generate sales. This presented a range of evidence noting that the partner's in-house location type classifications had limited representation of the times that stores generate sales. The statistical analysis of the clustering identified that a four-solution cluster group was the most appropriate solution. The representation of the clusters for the wider grocery store environment as indicators of temporal sales in stores was tested in chapter eight. In a sample of 136 stores, the four cluster groups were representative of a wider population 89% of the time. This evidence suggests that it is possible to group stores according to the cluster classification and that the patterns observed strongly represent actual temporal conditions and offer an effective representation of the partner's store network. Furthermore, it is then possible to isolate the typical uplifts in sales and draw comparisons with identifiable drivers of trade. The practical implications to retailers of this output are discussed further in section 10.4. However, from a temporal sales perspective the observed similarities within clusters and differences between clusters (and the temporal trends that appear to occur regularly at certain times of the day, in certain stores and in certain locations) indicate that similar temporal drivers of store level sales are occurring. Using this insight to understand the temporal drivers of trade it could be possible to predict temporal patterns of sales for stores and improve revenue estimation.

Analysis of the observed average temporal patterns for stores in each cluster group and store characteristics suggested that different demand types could be attributed to the uplifts in sales and that they were a major driving factor. Newing et al. (2013) note how in tourist areas of Cornwall, visitors are a major driver of seasonal sales at certain times of the year, supporting this conclusion. Comparisons were drawn suggesting that temporal peaks were likely related to the presence of specific demand types, which have considerable impact upon the volume and timings of store sales. For example, cluster 4 demonstrated that stores catering to considerable student populations were predominantly limited to sales in the evening as opposed to cluster one (workday convenience) which catered to considerable workplace populations and generated most sales at specific points during the working day. Interestingly, the 'student central' cluster was identifiable very early on in the cluster segmentation process, indicating an important distinction between the sales profiles of stores catering to students and stores catering to the general population. This is an important insight that retailers should consider as part of location planning and store management, because the evidence demonstrates that the prevalence of students can dramatically alter the trading profile of a store. The findings demonstrate novel insight into the times that stores make sales and improve our understanding of how sales fluctuate. The analysis of the data using clusters illustrates that though differences occur, groups of stores appear to follow distinguishable temporal patterns. Thus, patterns and subsequently total store revenue appears to be attributed to spatiotemporal demand (which receives ongoing discussion in section 10.3), suggesting that evidence of similar temporal activity is likely to be the result of the presence of certain demand types. The evidence strongly suggests that to accurately predict store revenue it is important to fully understand the spatiotemporal dynamics of catchment area demand and to represent spatiotemporal fluctuations in models as effectively as possible.

The analysis and outputs presented in chapter four address the current gap in the literature regarding present day temporal sales patterns, as well as presenting novel insight about the potential drivers of spatiotemporal store sales. The chapter has offered a deeper and novel understanding about the complex nature of the supply side in the UK grocery industry in relation to spatiotemporal store sales (grouping stores together that were previously in separate groups for instance) complimenting future academic study and analysis within commercial teams.

10.3 Summary of spatiotemporal demand

The second key aim of this thesis was to investigate the spatiotemporal geography of consumer demand. The outcome of this aim was to better understand how consumer demand varies during the day resulting from spatiotemporal fluctuations. Building on the evidence of spatiotemporal sales presented in chapter four the investigation of spatiotemporal demand geography was the focus of chapter five. Utilising the cluster groups produced in chapter four, it was possible to relate the spatiotemporal demand-side findings to the spatiotemporal supply-side patterns previously discussed. Analysis of spatiotemporal consumer geographies was conducted on observed consumer data; this took the form of a rich and extensive loyalty card dataset provided by the partner. The dataset consisted of over 29 million records of individual product transactions, representing over 900,000 unique loyalty card for stores in West Yorkshire over a 12-week period. The objective of the analysis of the loyalty card data was to demonstrate evidence of patterns in the observed consumer data of potential spatiotemporal behaviour, supporting the conclusions of the previous section (10.2) by drawing comparisons between the loyalty card data and the spatiotemporal analysis of store sales regarding temporal demand. Subsequently, from the insight provided by the observed spatiotemporal patterns of demand (in the loyalty card data) it was possible to generate spatiotemporal estimates of daytime and night time consumer populations for various demand types that indicated spatiotemporal patterns that correspond with the observed spatiotemporal sales profiles. The different demand types were selected following the analysis of temporal sales and were identified based on the timings of sales and which appeared to have an impact upon store revenue.

Access to commercial loyalty card data is rare in academia and thus, the insights into actual consumer behaviour in the grocery market represent novel developments in the literature (Newing et al., 2013a, Waddington et al., 2017). Subsequently, few studies provide an extensive analysis demonstrating the spatiotemporal geography of various demand types in relation to store level sales using actual loyalty card data. However, some studies have similarly called for more analysis of the spatiotemporal components of populations and have attempted to develop estimation techniques for considering fluctuations in population distributions over time (Martin et al., 2013, Martin et al., 2015). However, this research is not retail focused and their analysis does not make use of observed loyalty card distributions. Newing et al. (2013) and Newing et al. (2014) represent the few studies that consider the impact of spatiotemporal demand in relation to temporal store sales and revenue estimation using actual loyalty card data. The authors note that spatiotemporal demand remains an under-researched topic, and though primarily concerned with spatiotemporal affects of seasonal tourism demand, they highlight the academic and commercial need for continued spatiotemporal demand analysis in store revenue estimation, which this chapter addressed.

The total value of all loyalty card transactions for all stores in each cluster were aggregated up to LSOA level and then mapped. The maps (figures 5.1-5.8) demonstrate the origin of the sales and show the volume of sales and the count of unique customers for each cluster. The national distribution of customers and sales at stores indicate that sales are not always driven by residential demand. While the highest volume of sales and consumer counts are typically found in close proximity to the stores in each cluster (which is to be expected), the evidence of sales occurring beyond traditional catchments, and at national scales, suggest that spatiotemporal demand is occurring as those sales are unlikely to be the result of residential (local) based demand. To support the assumption that spatiotemporal demand was occurring, the core catchment for each cluster was calculated using the observed consumer flows from the loyalty card data. The technique adopted for this was based on common academic practice for catchment area analysis (Dolega et al., 2016), but using the partner's own analytical threshold for core catchment area identification, which uses the areal extent at which a total of 70% of total store revenue is met. Once this value is reached it was assumed that this area represents the extent of the core catchment. The output of this analysis was presented in table 5.2 and the results offered interesting insight into consumer behaviour with the evidence indicating patterns of spatiotemporal demand through the locality of loyalty card customer geographies.

It became apparent from the spatial pattern of loyalty card sales that many of the observed transactions repeatedly did not correspond with spatial patterns typically associated with local and residential shopping behaviour or with the characteristics of the stores in each cluster. For instance, cluster 1 indicated a catchment of 5km, which for a primarily convenience store based cluster group was extensive; there was also a relatively low residential population in the immediate vicinity of the stores, which further indicated this uncharacteristic consumer geography. In this instance it is unlikely that residents would travel the distances observed to a store of that type and size (c-stores) or one located in a city centre from their residential address to undertake their regular food shop. This is even less likely when consumers appear to have bypassed larger and closer stores. However, the illustrated spatial pattern of sales and volume of customers in the data suggests that sales are the result of regular journeys, such as journeys to work and that sales may occur when the store is in close proximity to a another location. Consequently, these transactions were likely the result of spatiotemporal demand originating from a separate location other than the consumer's home address, at different times of day, and would explain the temporal sales profiles observed in chapter four. Furthermore, in all clusters, there was evidence of sales originating from well outside of the core catchment area, in some instances the volume of these sales and count of individual customers were high, further suggesting regular shopping habits that could be the result of consumer demand moving around on a daily basis. Sales which occurred long distances from the stores appeared to be smaller in value and less regular (due to fewer customers) and could be associated with less frequent spatiotemporal dependant visits. Nonetheless, sales were observed and the evidence may be linked to tourists who have come to the area for a ‘one-off’ visit and purchased food during that period. Tourism demand has previously been demonstrated to have a considerable impact on store revenue and to be connected with fluctuating spatiotemporal consumer geographies (Newing et al., 2013a). Through an investigation of observed loyalty card data it has been possible to demonstrate that spatiotemporal demand is prevalent and is subsequently impacting grocery store revenue. The spatial patterns in the data appeared to suggest spatiotemporal activity for different demand types offering novel insight into temporal consumer behaviours. However, the loyalty card data (which was not time stamped) does not demonstrate the impact and level of change in consumer populations at store level geographies at a temporal scale.

To better understand the impact of spatiotemporal demand on store revenue and to illustrate the resulting geographies, a series of spatiotemporal demand estimates for day and night time periods were generated. The maps, which are shown in the latter stages of chapter five, demonstrate the spatiotemporal fluctuation for different demand types highlighting the regional and local transformations. Using a mixture of new data (WPZs) and a bottom-up approach using various census data, it was possible to create a series of novel demand layers for specific population groups such as, workers, school pupils, daytime visitors and university students for day and night times. The maps illustrated considerable fluctuations of people typically moving from their home addresses at night to a new location during the day. Flows appeared to concentrate around specific locations during the day time (e.g. schools, campus or visitor attractions) indicating high uplifts in potential consumers for stores in these locations at certain times. Conversely, though not unexpectedly, this resulted in outflows from ‘more-suburban’ areas resulting in fewer consumers in these locations during the day. The scale of spatiotemporal demand fluctuations was investigated at a cluster group core catchment level and was modelled for day and night time (shown in table 5.4 and 5.5). This supported the evidence presented in the demand maps, providing a detailed picture of spatiotemporal demand geographies in relation to grocery store demand. This evidence provided new insight that improves the understanding of spatiotemporal demand geographies and highlights why store revenue can vary throughout the day in response to spatiotemporal demand fluctuations. Furthermore, it demonstrates the importance of fully understanding spatiotemporal components of demand in store revenue estimation due to shifting patterns of expenditure and contributes new data, that if utilised, could satisfy the current need for store revenue estimation models to be better equipped for handling spatiotemporal components (Birkin et al., 2010, Hernandez, 2007, Hood et al., 2015, Newing et al., 2013a). The process of translating the insight generated from understanding spatiotemporal demand geography into a practical product, in relation to store revenue estimation and the subsequent outputs with implications for retailers are discussed in section 10.4.

While the analysis was insightful the methodology is not without limitation. For instance loyalty card data is only available for one retailer and it would therefore be unwise to assume a full representation of consumer profiles from the data of all and their own customers, which Birkin et al. (2010) highlight can result in data bias by demonstrating only the behaviours of the partner's customers. The spatiotemporal demand data were based on abstractions of census data and although freely available, which makes reproducibility more viable, the data itself is not temporally referenced which could result in spatiotemporal misrepresentation. However, to minimise the impact of this the disaggregation of demand by type and by time was applied logically and within the context of grocery shopping as discussed in chapter five. There are also commercial challenges that could impact upon the findings in a practical context. A current challenge experienced by many location planning teams is the struggle for internal legitimacy, particularly as many retailers decrease the level of store expansions and new investment programmes (Waddington et al., 2017). This could restrict the implementation of new data and analyses, as many departments shrink in size and focus more on the existing store network and already well established methods.

The evidence presented in chapter five contributes novel insights on spatiotemporal grocery demand behaviour, specifically in regard to understanding and demonstrating spatiotemporal geographies of different demand types and the observed impact to core catchment area demand at a store level. Furthermore, from a practical point of view it could be argued that through a better understanding of spatiotemporal demand patterns, retailers can take advantage of the insights regarding consumer behaviour and their locations at certain times of the day, thus resulting in operational changes to strategic decision making, affecting day-to-day operational concerns such as staffing, stocking, promotions and product lines at a store-by-store level.

10.4 Summary of models and forecasting

The third key aim of this thesis was to build on current SIM techniques, improving store revenue estimation by incorporating spatiotemporal components of demand, which represented an aspect of location planning that had not been fully explored. Initial development of a custom-built store revenue estimation model was discussed in chapter six. This took the form of a spatial interaction model (SIM), which was chosen because of it prevalent and continued use by retailers, particularly those in the grocery sector for store revenue estimation (Reynolds and Wood, 2010, Wood and Reynolds, 2011a). Adapting a well-established method helped to increase the traction of the outputs as they are directly comparable with current commercial procedures, and the results are typically well understood by management in retail, thus increasing the likelihood of positive and ongoing impacts. The effect of spatiotemporal components on the capacity of SIMs to estimate store revenue was then tested in chapter seven with the introduction of first a fully disaggregated workplace model, accounting for the spatiotemporal fluctuation of workplace consumers during the day. Previous evidence had demonstrated that during the day the population count of an area can change dramatically. This is especially the case in urban and city centre areas and sharp uplifts in sales that occur in stores in those areas have been shown to be related with transient workplace populations (Berry et al., 2016). Using a separate workplace population, chapter seven was able to demonstrate the impact of workplace based consumers on the grocery market using a SIM and demonstrated that revenue estimations were improved following the inclusion of a spatiotemporal and more realistic demand layer. Subsequently, the additional spatiotemporal demand types discussed in chapter five were then estimated and incorporated into a SIM in chapter eight.

In section 8.2, expenditure estimates and the typical grocery behaviour for each of the spatiotemporal demand types were generated. This was done using a series of logical, data driven assumptions, regarding both the estimation of individual demand expenditure and their grocery shopping behaviours, the spatiotemporal component having been previously demonstrated in chapter five. Individual expenditure estimates were generated for each demand type using both literature and commercial insight, This demand was then reallocated from a night time residential address to the temporary daytime address or added in at this location as new expenditure. As noted above demand was estimated using evidence from previous literature and commercial data to apply logical shopping behaviours to each of the demand types. For instance, the store sales data and loyalty card (LC) data appeared to indicate that university students typically have a low uptake in LC schemes due to the low number of LC customers located around the cluster four stores (as shown chapter five) and that when on campus they do not appear to travel far if purchasing food due to the low daytime sales (figure 8.23), which occur irrespective of the large campus population during the day which is less than 1km away (figure 8.24). This analysis offers advances to our understanding of grocery consumer demand, particularly for school and university based grocery consumers, filling the current gaps in the literature and commercial needs regarding the modelling of different spatiotemporal demand types (Berry et al., 2016, Newing et al., 2013a, Waddington et al., 2017). By addressing the redistribution of grocery expenditure that accompanies the daily activities of transient populations, through the novel revenue estimates and behaviour developments discussed in chapter eight, it was possible to create a series of disaggregated demand layers suitable for use in store revenue estimation. The different spatiotemporal components were then integrated into the final store revenue estimation model in Section 8.3; each demand type was calibrated separately in individual models and were then combined together to produce a final temporally influenced revenue estimation.

The development of the new SIM created greater accuracy of the revenue estimations. This was documented through chapters six and seven with the final revenue estimations presented in section 8.3.2. Although the number of revenue predictions within a +/-10% threshold totals only 8 stores, this represents a significant improvement on the night time only SIM, with an increased number of revenue estimations closer to the observed store revenues. Newing et al. (2014) noted that following their incorporation of seasonal tourism, revenue estimation was shown to be improved, and likewise it was also possible to observe the positive impact and incremental improvements to store revenue estimation at each stage of spatiotemporal extension documented in this thesis. From a theoretical perspective the spatiotemporal SIM was shown to be capable of handling different spatiotemporal demand types, in section 8.4, and demonstrated a reasonable capacity to distinguish the differences in demand behaviour making it possible to simulate various temporal scenarios that can affect store revenue. The refined daytime population estimates are far more realistic of actual consumer demand though there were limits to the accuracy of individual store estimates.

Previous literature has implied that current SIMs are less appropriate for micro spatial scale analysis, such as for predicting convenience store revenue (Birkin et al., 2002, Hood et al., 2015, Wood and Browne, 2007). The growth of this market (Harries, 2014, Hood et al., 2015, IGD, 2015) has resulted in considerable numbers of major grocery retailer's convenience stores and high numbers of symbols and independents, nationally, but also within the study area. It is likely that the complex nature of consumer interaction with convenience stores is difficult for a conventional SIM to handle, and as a consequence the convenience store revenue estimations were less impressive. However, the observed improvements demonstrated in the estimation of convenience store revenues highlight that through additional spatiotemporal components, location planning model predictions (e.g. SIMs) can be improved and made more capable of modelling the convenience store market. Anecdotally, Asda store location planning team noted that they do not include any smaller format stores in their current SIM to avoid this issue and suggested that this appeared to improve the quality of the supermarket store predictions, but that they also do not account for spatiotemporal demand. Another limitation of the SIM has been previously noted by Birkin et al. (2010). The authors comment on the issue of model boundaries, resulting in missing expenditure at stores located close to the edges of the study area chosen, which could account for a number of the under-predicting stores. They offer a solution in the way of boundary free modelling, whereby demand is not constrained by artificial boundaries, though they note that this is computationally intensive. There was also an increase in the number of stores over-predicting following the addition of spatiotemporal components. In some stores it was possible to link the over predictions with ethnic shopping behaviours (discussed in chapter two), for instance the Harehills area of Leeds has a high number of ethnic residents who are known to favour the increasing number of local and ethnic oriented stores. The store in this area was subsequently over-predicting, though future models can be improved by reducing volumes of demand spent in mainstream retailers, in areas with a high number of ethnic communities. It is likely additional factors are also important to consider, such as journey to work routes, which could account for under-predictions at some stores and over-predictions at others, as money is spent while on the way to work. Nevertheless, the aim was to demonstrate that store revenue estimation techniques and location based decision-making could be improved via spatiotemporal components and in this capacity this has been achieved. At the beginning of chapter eight, drive time data was introduced to improve the robustness and realism of the distance deterrence for consumers in the model, and while this is acknowledged in the literature to be better (Birkin et al., 2010), it presented a new challenge to overcome. The data, provided by the commercial partner used car travel times on the UK road network from LSOA to LSOA. This meant that the stores found within the origin LSOA were given an unrealistic drive time and made each of the stores overly attractive, as there was no travel cost (drive time) involved in the journey. This may have resulted in some of the over-predictions observed, however, to limit the impact in these instances additional time was added to the travel time to compensate for this data issue. It might be possible for ongoing research to improve the robustness of a minimum drive time threshold or to acquire different data that does not have the same issue. If future research is able to address the limitations, the spatiotemporal SIM will become even more robust, and I believe that the current observed improvements to accuracy would become even more apparent.

The findings of the long-term scenario modelling (section 9.2) regarding the impact of fluctuating student demand appear logical: the predictions represent a potential worst-case scenario, regarding the loss of sales when students are away on holiday on a store-by-store basis. For retailers, strategically it may be difficult to counteract the annual decline in sales experienced from fluctuating demand, such as students, especially when overall the store performance remains relatively consistent throughout the year and owing to the fact that they have little control over consumers' movements (and while the presence of one demand type may drop, the others remain present). Operationally, retailers may be able to run promotions at these points in the year when student numbers are down to increase the sales driven by other demand types as a means of balancing the potential revenue lost by fluctuating student demand. Additionally, this represents novel insight into student driven grocery sales and the impact of store sales that, as yet, has received limited attention in the relevant literature. Overall, the additional level of insight that a temporal approach in location planning can offer will likely be of benefit to retailers, that could potentially be missed when using a residential night time population. Retailers adopting this approach could obtain detail into the specific impact of the temporal components of individual demand types, which could then be modelled to represent different times/scales at will, as demonstrated in the long and short term scenarios.

Store revenue estimation was demonstrated in chapter eight to be improved through the utilisation of spatiotemporal demand. In section 9.3 the spatiotemporal model was applied in a set of location planning scenarios, for the estimation of potential new stores, to demonstrate the increased level of information that is gained from using spatiotemporal demand layers compared to a residential only model. In the new store revenue estimation scenarios, four store locations were modelled, three of which were based on recent store investments made by the retail partner, however the actual revenue data was not yet available. Each of the stores was also simulated in a residential only SIM and it was possible to compare the revenue estimations and predicted flows (£) in table 9.6 and figures 9.4 and 9.5. What is clear is that the core catchment area characteristics of each store are far more complex than indicated when represented by a residential population. Furthermore, the catchment areas appear to clearly demonstrate, though more obviously in the Vicar Lane and Bingley stores, an obvious spatiotemporal dynamic nature, demonstrated by a mixture of demand types present and absent at certain times of day. Thus, it appears that in estimating a new store's revenue, not only do location planning models that utilise spatiotemporal demand layers appear to be more representative of actual daytime populations (through the inclusion of transient demand types), but that in disaggregating demand types, models are capable of providing more detail. In the spatiotemporal model, the revenue estimations are higher than the residential only model. With the exception of the Royal Park Lane store, which appears to gain a major uplift in sales from the residential student populations (figure 9.3), the majority of the increase in sales for the other stores is driven by transient populations that arrive during the day, such as workplace and campus based consumers. Though the scenarios were a simulation, as previously noted, three of the stores represent real store developments made recently by the data partner. In these instances it was possible to demonstrate the importance of applying spatiotemporal demand for location planning decision-making. The differences in revenue estimations and breakdown of the temporal drivers of trade in the three stores (which were not identifiable in the residential only model), highlights the need to clearly understand the temporal components of core catchments, which vary throughout the daytime and appear to have a sizeable impact upon store revenue. The message is that the retail partner needs to consider the spatiotemporal components, because if left ignored there is a potential to overlook or underrepresent the profitable nature of a location, which is characterised by temporal demand types. Martin et al. (2015) noted that residential only populations were a poor representation of daytime populations and were typically characteristic of a night time population; the evidence in chapter five further highlighted this demonstrating the extent to which daytime population can vary in West Yorkshire. Through an example of a practical application of the SIM, section 9.3 demonstrated the lack of representation of a residential population, when compared to flows to stores using a spatiotemporal daytime population. Therefore, the evidence presented in section 9.3 demonstrates the need for location planners to adopt a series of spatiotemporal demand layers, evolving current techniques to include spatiotemporal components of demand, which will likely improve store revenue estimations. This can also improve the understanding of the demand-side drivers of trade and how variation at certain times of day will impact store revenue, at a store-by-store level. It is in the interest of the retail partner to test the spatiotemporal SIM developed in this thesis. Consequently, the researcher intends to request access to the average weekly sales data (once available) to compare the models revenue estimations of the stores modelled in section 9.3 with the revenue estimations.

An important novel development and potential practical application of the spatiotemporal analysis resulting from a non-traditional cluster segmentation of stores, was the use of the four temporal cluster profiles as a tool for predictive modelling. As well as providing novel contributions to academic research, helping to demonstrate the benefits of spatiotemporal data and improve our understanding of temporal sales fluctuations in stores, this also caters to the commercial needs of retailers, providing actionable insight with implications for store management at an operational level (Waddington et al., 2017). This represents a new revenue prediction technique, building on conventional analogue based modelling, which location planning teams will be able to take advantage of adding to the suite of already established tools. The evidence of regular and identifiable sales fluctuations in stores, which could be related to the presence and absence of specific spatiotemporal demand types meant that by understanding the catchment area characteristics of a store it appears possible to predict the sales profile by using the cluster groups, which was demonstrated in section 9.4. The methodology used was akin to the store revenue estimation analogue models already widely used by location planning teams (Reynolds and Wood, 2010, Wood and Reynolds, 2011a) using historical data to draw comparison between the test stores and the clusters.

To test this methodology, a totally different sample of stores was used and by relating each store's core catchment area characteristics with one of the cluster groups, it was possible to assign each store into one of the cluster profiles. Once segmented it was assumed that the average sales profile of the parent cluster was thus representative of the average sales profile to be expected at a store throughout a typical day. To test the temporal sales profile predictions a J48 decision tree process was used, which sorted the stores into the actual cluster group that each store should belong to according to their observed sales data. The output and accuracy of the predictions indicated that it was possible to estimate the type of sales profile to be expected for stores correctly, 76% of the time. The potential foresight regarding a new store provides retailers with unique opportunities for operational decision-making in stores that until now were typically based on data collected following a short review period after opening. However, as demonstrated, it appears possible to predict the type of sales profile to expect in a store allowing retailers to optimise store operations according to the sales profile of the cluster in which a new store is allocated. Hence, management can adapt staff rotas, adjust a store's stock and promotions in response to temporal fluctuations or dynamically adapt promotions at different times of day depending on the sales profile and demand types present or indeed absent at that time. The key argument (and major implication for location planning) is that this methodology allows retailers to make these decisions in advance of the store opening; retailers are able to open stores optimised to their catchments because they are able to make use of crucial temporal insight about supply and demand (resulting from this analysis) during the location planning stage of store development.

In summary, this analysis offers novel insight and a new tool that has not been previously demonstrated in the literature and it offers location planners insight which could boost store efficiency and performance due to stores opening ‘optimised’ for their catchment area. Furthermore, the early evidence presented in chapter 8 and section 9.3, demonstrate the increases in accuracy and additional level of detail regarding spatiotemporal demand and grocery store revenue estimation that was achieved in this thesis from the inclusion of temporal components. Moreover, the research has satisfied the need, noted by Newing et al. (2013) and Waddington et al. (2017), for increased academic attention into spatiotemporal components in regard to store location planning by presenting novel analyses and a better understanding of the impact of spatiotemporal demand on current grocery store revenue estimation techniques.

10.5 Summary of practical benefits

The bullet points (below), which are a summary of the practical benefits on offer as a result of this research, have been addressed throughout this work. However, this section briefly summarises the potential practical implications and benefits that the analysis presented in this thesis offers for location planning decision-making:

- Improved representation of the distribution of daytime populations.
- A better range of grocery shopping behaviour through the increased disaggregation and use of novel demand types.
- Estimation of convenience store revenue appears to improve when using a daytime demand model, compared with a residential only model. The potential to overcome issue surrounding convenience store revenue estimation techniques using conventional SIMs, which typically have limited success (Wood and Browne, 2007), for instance, due to the nature of a (non-traditional) footfall driven or residential based catchments, has been addressed, demonstrating that revenue estimations can be improved through a more realistic representation of localised and temporal catchment demand surrounding convenience store sites.

- Demonstrated flexibility in the SIM approach, which can be updated as more and better spatiotemporal data becomes available regarding demand with the potential to improve the accuracy of temporal representation in the future.
- New store segmentation type (using temporal sales) with operational implications for store development and management.
- Ability to estimate the temporal sales profile of stores for an average day of sales.
- In response to the predicted sales profile, which offer informed decision-making, the ability to plan staffing, restocking and the type of products on offer by time of day.
- Novel insight into the dynamics and temporal nature of catchments at a store level through a better understanding of catchment demand.
- Insights into the specific impact that temporal/transient population types can have on store revenue.

10.6 Thesis contribution

The opportunity to work with a commercial partner and utilise their vast and recent data sets helped the applied nature of this research. By working closely with the commercial partner it was possible to identify current areas of weakness (relating to their SIM location planning methods) and to address aspects of research that due to limited resources were (at present) not a priority, even though they presented an area of interest. The research presented in this thesis has been presented at several engagement events held by the Consumer Data Research Centre (CDRC) in the University of Leeds. These events were aimed at integrating academic research with the commercial sector and to showcase the types of analysis academic research can offer. The applied focus and close relationship with the data partner has meant that this research has often been at the forefront of these events, highlighting the academic and commercial benefits to industry resulting from a symbiotic relationship (that enable both academic and commercial stakeholders to benefit and develop novel research) and the research presented has been well-received and generated interest from commercial partners. There have also been regular talks with the data partner, where it has been possible to convey the research agenda in more detail, ensuring that the novel analysis has addressed their interests and contributed to their location planning methods.

The thesis has made use of rarely available industry data and while addressing areas of genuine commercial interest, the novelty and rarity of the data has meant that the research has addressed several of the main academic needs (relating to SIM location planning) and under-researched areas within the academic literature. For instance, writing in 2007, (Wood and Browne, 2007) suggested that location planners are facing difficulties in estimating convenience store sales, which are affected by highly localised catchment areas. This research has demonstrated that through an increased analysis of catchment areas and representation of spatiotemporal demand it is possible to improve the performance of SIMs (chapters seven and eight) when applied to convenience store modelling. Moreover, the major outputs from chapters four and five, which have been disseminated via a publication (detailed below), has addressed the need for an increased focus on spatiotemporal demand (owing to evidence of the impact of spatiotemporal demand on store sales) especially with regards to location planning and the lack of representation of temporal consumer populations (Berry et al., 2016, Newing et al., 2013a, Newing et al., 2014b). The publication presents a novel segmentation of grocery stores according to four temporal sales profiles and relates temporal fluctuations in sales with specific spatiotemporal demand-side drivers of trade. Furthermore, the research builds on two decades of analysis, presenting a recent review of the diurnal activity of consumers shopping in supermarket stores providing context on present day consumer behaviours.

Waddington, T.B.P., Clarke, G.P., Clarke, M. and Newing, A. 2017 Open all hours: Spatiotemporal fluctuations in UK grocery stores and catchment area demand. International Journal of Retailing, Distribution and Consumer Research, 1-26.

At the time of submission a second paper originating from this thesis (specifically focusing on the work demonstrated in section 9.4, regarding the prediction of temporal sales profiles of stores) was planned, and it is hoped that this will provide a practical application for location planners to add to their established modelling techniques. The proposed paper will seek to demonstrate the value of the cluster analysis, which was published in the first paper, in an applied context highlighting the practical benefits and methodologies developed in this thesis, as well as building on the (surprisingly under-researched) store location planning literature regarding temporal sales analysis. This thesis and its associated publication provide a clear contribution to the academic literature, and are related to industry practice throughout, demonstrating the practical need and justification of the applied focus and research agenda.

10.7 Limitations and recommendations for future research

In order to ensure increased robustness and future proofing of the outputs presented throughout this work, there is a need for continued research. Below, are the recommendations for future research that I believe are important steps to take as part of the ongoing location modelling literature and commercial practice relating to this theses' agenda:

One of the limiting factors of the SIM was the data, while the dataset provided by the retail partner was insightful and was relatively recent, much of the demand-side data was generated using census based data. The problem with census based data is that it represents a static snapshot of the population at a specific point in time and due to the 10 year cycle can potentially be inaccurate after release. Therefore, the first recommendation for future research is to further improve the quality and detail of the spatiotemporal demand data, which is used to derive the origins of different demand types at certain times of day. At present the data used to derive the spatiotemporal demand layers for a daytime consumer distribution was collated from a series of census and workplace statistic based data. While it was possible to estimate the volume and location of different demand types during the daytime, as more and better temporal data which is increasingly being collected by commercial businesses and made available to academia, it is likely that improvements to the representation of the distribution and volume of consumers both spatially and temporally, through new datasets, will offer increased accuracy for modelling consumers throughout the day. The utilisation of these data represents an important action for continued research to secure, which will be particularly useful for improving the estimation of available expenditure within a store catchment area at certain times of day, as well as for commercial location planning teams to ensure that location based decisions are informed with up-to-date and accurate demand layers. For instance, it is becoming increasingly possible to utilise the spatiotemporal insight generated from data collected by social media websites or mobile phone companies, such as Twitter or Telefonica (the owner of O_2), for academic research and which have already received some attention in recent years regarding the modelling of diurnal population movements, demonstrating this potential (Birkin et al., 2017, Birkin et al., 2014, Lovelace et al., 2016, Malleson and Birkin, 2013a, Malleson and Birkin, 2013b). An alternative source of spatiotemporal population movements and counts of consumers could be derived by using the emerging data collected by Local Data Company (LDC). LDC use mobile phone signals (for phones which are searching for a Wi-Fi signal), which pass by a series of physical sensors, typically placed in retail units. The sensors record the number of individuals that pass and the time of day that this occurs, and unlike the mobile phone data provided by Telefonica, their data is not restricted to one mobile provider, though they are limited by the number and spatial location of monitoring devices and to phones that are trying to connect to a Wi-Fi signal. However, in all of the examples it is possible to derive the volume of individuals in any one area at anyone time, with a varying degree of accuracy and robustness. These types of data are not without their own limitations, but have shown promise and could therefore be used to create accurate accounts of the diurnal distribution of consumers. Furthermore, through spatial and temporal analysis it would be possible to infer the demand type based on where and when they are from the data. For example, Malleson and Birkin (2013, 2014) were able to gain insight into individuals and their behaviour via a content analysis of users tweets, in addition to the time and location of the tweet. This could be used to generate a series of disaggregated spatiotemporal grocery shopping demand types such as those demonstrated in this research.

The second series of recommendations for ongoing research refer to the SIMs. It would be useful to recreate the residential only and the spatiotemporal (daytime) demand SIMs for a different region. While West Yorkshire is a representative region of the average UK grocery consumer and thus the SIM is likely to be a fair representation of typical consumers anywhere in the UK. It is a potential limitation of this research that the SIM was only built for West Yorkshire. Therefore, it would be beneficial for future research to recreate the model for a new area so that it is possible to test the robustness of the spatiotemporal demand layers and to review the benefits of spatiotemporal demand disaggregation on store revenue estimation. This would not be a difficult task for ongoing research, securing commercial data regarding store revenue and loyalty card consumer behaviour due to the good working relationship with the commercial data partner. This continued relationship and opportunity to work with recent and observed commercial data will further strengthen the analyses and robustness of the SIMs, but it also presents opportunities to expand the research nationally, due to the partners national level store network. This research has discussed the lack of suitability of conventional SIMs for the planning and revenue estimation of convenience stores. While chapter three has shown that SIMs are an important and common location planning tool in the grocery industry, their use for smaller store formats has been limited. This has been attributed to their inability to capture and recreate the local and small-area drivers of trade that occur in convenience stores, such as the impact of footfall driven by transient populations. However, this research has demonstrated that with increased spatiotemporal disaggregation (generating demand layers that are more representative of daytime populations and accounting for the movements of consumers at certain times of day) SIM accuracy for convenience store revenue estimation can be improved, when compared to a residential only model. Nevertheless, it is suggested that further work seek to apply small-area spatiotemporal data with greater detail and build upon the expenditure estimates developed in this thesis for transient demand types for the benefit of convenience store revenue estimation, to further improve reliability of the SIM for small store format location planning through a more detailed portrayal of localised footfall using the potential data sets discussed above.

A third recommendation would be to continue work on the temporal clusters and the estimation of the diurnal sales profiles developed in chapters four and nine. It is suggested that further work seek to expand the number of stores used to derive the cluster groups. The initial sample size could potentially be a limiting factor effecting the representation of the clusters for the wider store network. The use of a larger sample of stores in the K-means cluster analysis could generate additional temporal sales profiles for stores, by analysing areas of the UK that are not represented in the West Yorkshire dataset, and further increasing original cluster representation from 87%. For instance, the work of Newing et al. (2013) found that for stores located in coastal resort tourist destinations, store revenue was impacted greatly by the presence of tourists. It is likely therefore that in using a larger store sample to generate the cluster groups, which would then be used to estimate diurnal sales patterns in new stores (as in chapter nine), the increased sample will be more representative of the national store sales profiles. Similarly, this may result in the identification of additional sales profiles for stores with distinct diurnal sales profiles, such as the student central cluster group improving the accuracy of the diurnal sales profile analogue model and our understanding of fluctuating temporal sales and the drivers of the fluctuations. Furthermore, future research should seek to refine the cluster group membership criteria, building a series of more robust rules so that location planners when drawing on analogies are able to more robustly state which cluster a new store development would likely belong to. This could be aided by the use of time-stamped loyalty card data; this would make it possible to relate consumer mobility patterns at certain times of day to specific demand types as already demonstrated, but additionally it would be possible to relate the peaks in store sales profiles with the purchasing of certain types of products and certain demand types by time of day. This could reveal interesting insights into consumer behaviour throughout the day, but could also provide further detail in the segmentation of stores, by temporal sales profiles, making it possible to relate the presence of a specific demand type and their impact on store performance and store operations throughout the day (by the products they typically purchased, aiding the location planning and operational decision making).

In summary, it is important for future research to address the limitations discussed through this thesis and to validate the models and the spatiotemporal data that are used to inform them. Furthermore, it is crucial that future research continues the dialogue and builds on the relationship between academics and commercial representatives. This research has shown that each side have a lot to offer and gain from working together. The opportunity to work with commercial data was not only a rare opportunity that resulted in novel developments for the discipline of location planning literature, but it was a major factor influencing the applied nature of this research. Subsequently, it is suggested that future research continue to feedback research outputs to the data partner, as well as wider commercial and academic audiences, such as through the contributions of the novel analyses presented in this thesis and in literature (Waddington et al., 2017). Given the important role of location planning teams, they are commonly centrally located within the corporate structure of grocery retailers (Wood and Reynolds, 2011b), therefore there is potential for the insights presented in this thesis (via a continued dialogue) to influence the decision-making process. Furthermore, a continued relationship will help to secure the development for ongoing and applied research from academia, through access to detailed and up-to-date commercial data, helping to ensure that research remains accurate, timely and future proofing analyses.

10.8 Concluding remarks

This research has quantified the temporal trends in store sales and developed our understanding of the patterns of temporal trade and the demand-side drivers of temporal fluctuation at a store level. Furthermore, this research has developed our understanding of the related spatiotemporal demand-side fluctuations of consumers, particularly the impact of transient consumer populations on store sales. It is hoped that through the development of SIMs, used to support location-planning decision-making, the commercial application of this research will have an ongoing and direct impact on location planning teams in the grocery sector with novel developments that build on established methods. From the perspective of an (early career) academic researcher the opportunity to work on research with both commercial and academic impacts with opportunities for future research agenda has provided a challenging, yet exciting opportunity. There are, of course, continued opportunities and a need for ongoing research, refining methods and increasing the robustness of revenue estimations, but this research with the current scope of available data has demonstrated the need and capabilities of location-based decision making supported by spatiotemporal data, that will continue to get better, as more and better quality data becomes available in both the commercial sector and for academic researchers. This study has addressed the aims of this research and offers timely analyses and novel outcomes that address the academic needs laid out throughout this thesis. Writing in 2012, (Ince and Jackson, 2012) noted that the opportunity for development and innovation from the understanding of consumers was crucial for retailers, and that this is increasingly being facilitated by academic research. Subsequently, it is hoped that this research offers an example of the successful collaboration between academic researchers and commercial practitioners, leading to advances in our understanding of consumer behaviour and the practical implications for retailers, as well as fulfilling the requirements of an academic thesis.

11.0 List of References

Level the playing field: The market [Online]. http://leveltheplayingfield.co.uk/the-market/. [Accessed 7.6. 2015].

ADMIN. 2016. *How our working habits will change towards 2020* [Online]. Service Futures: ISS World. Available: http://servicefutures.com/workplace-management/working-habits-will-change-towards-2020/ [Accessed 11/8/2016 2016].

AGGARWAL, C. 2013. An Introduction to Cluster Analysis. *Data Clustering.* Chapman and Hall/CRC.

AKEHURST, G. 1984. 'Checkout': The Analysis of Oligopolistic Behaviour in the UK Grocery Retail Market. *The Service Industries Journal,* 4**,** 189-242.

ALCOHOL CONCERN 2011. Out of the way? Alcohol displays in supermarkets. *In:* ALCOHOL CONCERN CYMRU (ed.). Cardiff.

ARENTZE, T. & TIMMERMANS, H. 2000. Albatross: a learning based transportation oriented simulation system. *The Hague: European Insitute of Retailing and Service Studies*.

ARENTZE, T. A., OPPEWAL, H. & TIMMERMANS, H. J. P. 2005. A Multipurpose Shopping Trip Model to Assess Retail Agglomeration Effects. *Journal of Marketing Research,* 42**,** 109-115.

AU, N. & LAW, R. 2002. Categorical classification of tourism dining. *Annals of Tourism Research,* 29**,** 819-833.

AVERY, J., KOZINETS, R., MITTAL, B., RAGHUBIR, P. & WOODSIDE, A. G. 2013. *Consumer behavior: human pursuit of happiness in the world of goods,* Cincinnati, Ohio, Open Mentis.

BAHN, D. L. & FISCHER, P. P. 2003. Clicks and mortar: balancing brick and mortar business strategy and operations with auxiliary electronic commerce. *Information Technology and Management,* 4**,** 319-334.

BAKER, R. G. 1994. An Assessment of the Space‐Time Differential Model for Aggregate Trip Behavior to Planned Suburban Shopping Centers. *Geographical Analysis,* 26**,** 341-363.

BAKER, R. G. 1996. Multi-purpose shopping behaviour at planned suburban shopping centres: a space-time analysis. *Environment and Planning A,* 28**,** 611-630.

BAKOS, Y. 2001. The Emerging Landscape for Retail E-Commerce. *The Journal of Economic Perspectives,* 15**,** 69-80.

BARON, S., HARRIS, K., LEAVER, D. & OLDFIELD, B. M. 2001. Beyond convenience: the future for independent food and grocery retailers in the UK. *The International Review of Retail, Distribution and Consumer Research,* 11**,** 395-414.

BATTY, M. & MACKIE, S. 1972. The calibration of gravity, entropy, and related models of spatial interaction. *Environment and Planning,* 4**,** 205-233.

BBC. 2015. *Global Recession Timeline* [Online]. BBC News Business: BBC News. Available: http://news.bbc.co.uk/1/hi/business/8242825.stm [Accessed July 20. 2015].

BEATTY, S. E., GIVAN, A. M., FRANKE, G. R. & REYNOLDS, K. E. 2015. Social Store Identity and Adolescent Females' Store Attitudes and Behaviors. *Journal of Marketing Theory and Practice,* 23**,** 38-56.

BELL, D. R. & LATTIN, J. M. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Science,* 17**,** 66-88.

BELL, M. 2015. Demography, time and space. *Journal of Population Research,* 32**,** 173-186.

BENOIT, D. & CLARKE, G. 1997. Assessing GIS for retail location planning. *Journal of retailing and consumer services,* 4**,** 239-258.

BERRY, T., NEWING, A., BRANCH, K. & DAVIES, D. 2016. Using workplace population statistics to understand retail store performance *The International Review of Retail, Distribution and Consumer Research***,** 1-21.

BHARGAVA, N., SHARMA, G., BHARGAVA, R. & MATHURIA, M. 2013. Decision tree analysis on j48 algorithm for data mining. *Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering,* 3**,** 1114-119.

BIRKIN, M., CLARKE, G. P. & CLARKE, M. 2002. *Retail Geography and intelligent network planning,* Chichester, Wiley.

BIRKIN, M., CLARKE, G. P. & CLARKE, M. 2010. Refining and Operationalizing Entropy-Maximizing Models for Business Applications. *Geographical Analysis,* 42**,** 422-445.

BIRKIN, M., CLARKE, G. P. & CLARKE, M. 2017. *Retail location planning in an era of multi-channel growth,* Oxon, Routledge.

BIRKIN, M., CLARKE, G. P., CLARKE, M. & WILSON, A. 1996. *Intelligent GIS: location decisions and strategic planning,* Cambridge, GeoInformation International.

BIRKIN, M., HARLAND, K. & MALLESON, N. 2013. The classification of space-time behaviour patterns in a British city from crowd-sourced data. *Computational Science and Its Applications–ICCSA 2013***,** 179-192.

BIRKIN, M., HARLAND, K., MALLESON, N., CROSS, P. & CLARKE, M. 2014. An Examination of Personal Mobility Patterns in Space and Time Using Twitter. *International Journal of Agricultural and Environmental Information Systems (IJAEIS),* 5**,** 55-72.

BIRKIN, M. & MALLESON, N. 2013. Investigating the Behaviour of Twitter Users to Construct an Individual-Level Model of Metropolitan Dynamics. *Working paper.* National Centre for Research methods.

BLYTHE, J. 2013. *Consumer behaviour,* London, SAGE.

BOWLBY, S. R. 1979. *Accessibility, Mobility, and Shopping Provision,* Oxford, Pergamon Press Ltd.

BRIGGS, F. 2010. Spar UK Clinches £150m Holiday park contract. Retail News: Retailer News.

BROWN, J. R. & DANT, R. P. 2014. The Role of e-Commerce in Multi-Channel Marketing Strategy. *Handbook of Strategic E-Business Management***,** 467-487.

BRYNJOLFSSON, E., HU, Y. & RAHMAN, M. S. 2009. Battle of the Retail Channels: How Product Selection and Geography Drive Cross-Channel Competition. *Management Science,* 55**,** 1755-1765.

BURT, S. & SPARKS, L. 1994. Structural change in grocery retailing in Great Britain: a discount reorientation? *International Review of Retail, Distribution and Consumer Research,* 4**,** 195-217.

BURT, S. & SPARKS, L. 2003a. E-commerce and the retail process: a review. *Journal of Retailing and Consumer Services,* 10**,** 275-286.

BURT, S. L. & SPARKS, L. 2003b. Power and Competition in the UK Retail Grocery Market. *British Journal of Management,* 14**,** 237-254.

CARAHER, M., LLOYD, S. & MADELIN, T. 2014. The "School Foodshed": schools and fast-food outlets in a London borough. *British Food Journal,* 116**,** 472-493.

CAVILL, N. & RUTTER, H. 2013. Obesity and the environment: regulating the growth of fast food outlets. *In:* PUBLIC HEALTH ENGLAND (ed.). Department of Health: Local Government Association (LGA), Chartered Institute of Environmental Health (CIEH).

CHARLES-EDWARDS, E. & BELL, M. 2013. Estimating the service population of a large metropolitan university campus. *Applied Spatial Analysis and Policy,* 6**,** 209-228.

CHENG, E. W., LI, H. & YU, L. 2007. A GIS approach to shopping mall location selection. *Building and Environment,* 42**,** 884-892.

CLARKE, I. 2000. Retail power, competition and local consumer choice in the UK grocery sector. *European Journal of Marketing,* 34**,** 975-1002.

COOMBES, M. 2010. Defining Labour Market Areas by Analysing Commuting Data: Innovative Methods in the 2007. *Technologies for Migration and Commuting Analysis: Spatial Interaction Data Applications: Spatial Interaction Data Applications***,** 227.

CUNNINGHAM, C. J. L., WEATHINGTON, B. L. & PITTENGER, D. J. 2013. *Understanding and conducting research in the health sciences,* Hoboken, New Jersey, Wiley.

DASHWOOD, R. 2013a. Robin Peston Goes Shopping: Episode 1. *Robin Peston Goes Shopping*.

DASHWOOD, R. 2013b. Robin Peston Goes Shopping: Episode 3. *Robin Peston Goes Shopping*.

DAWSON, J. 1980. *Retail Geography (RLE Retailing and Distribution),* Routledge.

DAWSON, J. 2000. Retailing at century end: some challenges for management and research. *The International Review of Retail, Distribution and Consumer Research,* 10**,** 119-148.

DE KERVENOAEL, R., HALLSWORTH, A. & CLARKE, I. 2006. Macro-level change and micro level effects: A twenty-year perspective on changing grocery shopping behaviour in Britain. *Journal of Retailing and Consumer Services,* 13**,** 381-392.

DEMOULIN, N. T. & ZIDDA, P. 2008. On the impact of loyalty cards on store loyalty: Does the customers' satisfaction with the reward scheme matter? *Journal of Retailing and Consumer Services,* 15**,** 386-398.

DENNETT, A. R. 2010. *Understanding Internal Migration in Britian at the start of the 21st Century.* Doctor of Philosophy University of Leeds.

DEPARTMENT OF EDUCATION 2015. Schools, Pupils and their characteristics. 16/07/15 ed. http://www.gov.uk: Department of Education.

DEPARTMENT OF THE ENVIRONMENT 1996. Planning Policy Guidance 6 (Revised): Town Centres and Retail Developments. London: Department of the Environment.

DESAI, K. K. & TALUKDAR, D. 2003. Relationship between product groups' price perceptions, shopper's basket size, and grocery store's overall store price image. *Psychology and Marketing,* 20**,** 903-933.

DESHMUKH, S. P., DESHMUKH, P. & THAMPI, G. 2013. Transformation from E-commerce to M-commerce in Indian Context. *International Journal of Computer Science Issues (IJCSI),* 10**,** 55-60.

DEVINE, P., LLOYD, K. & GRAY, A. 2006. University Student Food Attitudes and Behaviour Survey Final Report. *Belfast: Northern Ireland Social and Political Archive (ARK)*.

DIPLOCK, G. & OPENSHAW, S. 1996. Using simple genetic algorithms to calibrate spatial interaction models. *Geographical Analysis,* 28**,** 262-279.

DOLEGA, L., PAVLIS, M. & SINGLETON, A. 2016. Estimating attractiveness, hierarchy and catchment area extents for a national set of retail centre agglomerations. *Journal of Retailing and Consumer Services,* 28**,** 78-90.

DUDDING, V. & RYAN, C. 2000. The impacts of tourism on a rural retail sector: a New Zealand case study. *Tourism Economics,* 6**,** 301-319.

DUKE, R. C. 1989. A structural analysis of the UK grocery retail market. *British Food Journal,* 91**,** 17-22.

EAST, R., LOMAX, W., WILLSON, G. & HARRIS, P. 1994. Decision Making and Habit in Shopping Times. *European Journal of Marketing,* 28**,** 56-71.

EINAV, L., LEVIN, J., POPOV, I. & SUNDARESAN, N. 2014. Growth, Adoption, and Use of Mobile E-Commerce. *The American Economic Review,* 104**,** 489-494.

ELLIOTT, S., TWYNAM, B. & CONNELL, S. 2012. Building for breakthroughs. *The Korn/Ferry Institute*.

EUROMONITOR INTERNATIONAL. 2015. *Alcoholic Drink in the United Kingdom* [Online]. http://www.euromonitor.com/alcoholic-drinks-in-the-united-kingdom/report. [Accessed 7.6. 2015].

EVERITT, B., LANDAU, S. & LEESE, M. 2001. *Cluster analysis,* New York;London;, Arnold.

FALK, J. 2014. *Physical Mobile Interaction in Omni-Channel Retailing: Using the customers' smartphone to interact with smart objects in a store.* Linköping University.

FERNIE, J., FERNIE, J., SPARKS, L. & MCKINNON, A. C. 2010. Retail logistics in the UK: past, present and future. *International Journal of Retail & Distribution Management,* 38**,** 894-914.

FNS 2007. Consumer attitudes to food standards: wave 7. *In:* FOOD STANARD AGENCY (ed.). London: TNS.

FOTHERINGHAM, A. S. 1983. Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. *Environment and Planning A,* 15**,** 1121-1132.

FOTHERINGHAM, A. S. 1984. Spatial flows and spatial patterns. *Environment and Planning A,* 16**,** 529-543.

FOTHERINGHAM, A. S. 1986. Modelling hierarchical destination choice. *Environment and Planning A,* 18**,** 401-418.

GAURI, D. K., SUDHIR, K. & TALUKDAR, D. 2008. The Temporal and Spatial Dimensions of Price Search: Insights from Matching Household Survey and Purchase Data. *Journal of Marketing Research,* 45**,** 226-240.

GENTINA, É. & CHANDON, J.-L. 2013. Adolescent shopping behaviour: Different assimilation and individuation needs in France and the United States. *Journal of Retailing and Consumer Services,* 20**,** 609-616.

GINLEY, M. U.S. Beverage Alcohol Trends. U.S. Drinks Conference, 2013 Orlando, FL.

GOLDMAN, A. & HINO, H. 2005. Supermarkets vs. traditional retail stores: diagnosing the barriers to supermarkets' market share growth in an ethnic minority community. *Journal of Retailing and Consumer Services,* 12**,** 273-284.

GROEPPEL-KLEIN, A., THELEN, E. & ANTRETTER, C. 1999. The impact of shopping motives on store assessment. *European Advances in Consumer Research,* 4**,** 63-72.

GROVE, J. 2015. Vehicle licensing statistics: 2014 Quarter 4. *In:* TRANSPORT, D. F. (ed.). Department for Transport: GOV.UK.

GUY, C. 1991. Spatial interaction modelling in retail planning practice: the need for robust statistical methods. *Environment and Planning B: Planning and Design,* 18**,** 191-203.

GUY, C. M. 1996. Grocery store saturation in the UK – the continuing debate. *International Journal of Retail & Distribution Management,* 24**,** 3-10.

HALLSWORTH, A. & BELL, J. 2003. Retail change and the United Kingdom Co-operative Movement-new opportunity beckoning? *The International Review of Retail, Distribution and Consumer Research,* 13**,** 301-315.

HALLSWORTH, A., DE KERVENOAEL, R., ELMS, J. & CANNING, C. 2010. The food superstore revolution: changing times, changing research agendas in the UK. *The International Review of Retail, Distribution and Consumer Research,* 20**,** 135-146.

HAMLETT, J., BAILEY, A. R., ALEXANDER, A. & SHAW, G. 2008. Ethnicity and Consumption South Asian food shopping patterns in Britain, 1947—75 1. *Journal of Consumer Culture,* 8**,** 91-116.

HANGHANG, T. & KANG, U. 2013. Big Data Clustering. *Data Clustering.* Chapman and Hall/CRC.

HARLAND, K. 2008. *Journey to Learn: Geographical Mobility and Education Provision.* Doctor of Philosophy, University of Leeds.

HARRIES, J. 2014. Growing UK convenience market. The Institute of Grocery Distribution.

HARRIS, L. C. & OGBONNA, E. 2001. Competitive advantage in the UK food retailing sector: past, present and future. *Journal of Retailing and Consumer Services,* 8**,** 157-173.

HAVEMANN, J. 2015. *The Financial Crisis of 2008: Year in Review 2008* [Online]. Encycolpaedia Britannica Online: Encycolpaedia Britannica. Available: http://www.britannica.com/topic/Financial-Crisis-of-2008-The-1484264 [Accessed July 20. 2015].

HEITZ-SPAHN, S. 2013. Cross-channel free-riding consumer behavior in a multichannel environment: An investigation of shopping motives, sociodemographics and product categories. *Journal of Retailing and Consumer Services,* 20**,** 570-578.

HENRY, V. 2015. The rise and rise of online grocery shopping. The Institute of Grocery Distribution

HERNANDEZ, T. 2007. Enhancing retail location decision support: The development and application of geovisualization. *Journal of Retailing and Consumer Services,* 14**,** 249-258.

HILLIER, A., CANNUSCIO, C. C., KARPYN, A., MCLAUGHLIN, J., CHILTON, M. & GLANZ, K. 2011. How far do low-income parents travel to shop for food? Empirical evidence from two urban neighborhoods. *Urban Geography,* 32**,** 712-729.

HOOD, N., CLARKE, M. & CLARKE, G. P. 2015. Segmenting the growing UK convenience grocery market for store location planning. *International Journal of Retailing, Distribution and Consumer Research,* 26**,** 113-136.

HORNIK, J. 1984. Subjective vs. objective time measures: A note on the perception of time in consumer behavior. *Journal of Consumer Research,* 11**,** 615-618.

HUBBARD, P. 2009. Geographies of studentification and purpose-built student accommodation: leading separate lives? *Environment and planning A,* 41**,** 1903-1923.

HUFF, D. L. 1963. A probabilistic analysis of shopping center trade areas. *Land Economics,* 39**,** 81-90.

HUGHES, R., HALLSWORTH, A. G. & CLARKE, G. P. 2009. Testing the effectiveness of the proposed UK 'competition test'. *The Service Industries Journal,* 29**,** 569-590.

HUMBY, C. & HUNT, T. 2003. *Scoring points: How Tesco is winning customer loyalty (summary),* Kogan Page Publishers.

HUNTER, J. 2004. Grocery is growing. *Packaging Magazine* Tonbridge: Haymarket Publishing LTD.

IGD 2012. Shopper Trends. The Institute of Grocery Distribution

IGD 2014. Growing UK convenience market. The Institute of Grocery Distribution.

IGD 2015. Convenience Retailing Factsheet. The Institute of Grocery Distribution.

INCE, M. & JACKSON, B. 2012. Retail Revolution *Society Now.* ESRC ESRC.

INTERNATIONAL OFFICE UOE. 2016. Food shopping guide. *Estimated living costs 2016-17* [Online]. Available: http://www.ed.ac.uk/studying/international/finance/cost-of-living.

IPSOS-RSL 2003. United Kingdom Time Use Survey, 2000, [data collection]. *UK Data Service.* Office of National Statistics.

JETTER, K. M. & CASSADY, D. L. 2006. The availability and cost of healthier food alternatives. *American Journal of Preventive Medicine,* 30**,** 38-44.

JOCHEM, W. C., SIMS, K., BRIGHT, E. A., URBAN, M. L., ROSE, A. N., COLEMAN, P. R. & BHADURI, B. L. 2013. Estimating traveler populations at airport and cruise terminals for population distribution and dynamics. *Natural Hazards,* 68**,** 1325-1342.

JONES, P. M. B. 1982. Hypermarkets Superstores-Saturation or future growth? *Retail and Distribution Management,* 10**,** 20-27.

KAMARULZAMAN, Y. 2010. Geodemographics of travel e-shoppers: An empirical analysis of UK consumers. *European Journal of Social Sciences,* 16**,** 195-205.

KANTAR-WORLDPANEL. 2015. *Great Britain: Grocery Market Share* [Online]. London: Kantar Group. Available: http://www.kantarworldpanel.com/en/grocery-market-share/great-britain [Accessed].

KANTAR-WORLDPANEL. 2017. *Great Britain: Grocery Market Share* [Online]. London: Kantar Group. Available: http://www.kantarworldpanel.com/en/grocery-market-share/great-britain [Accessed].

KEH, H. T. & PARK, S. Y. 1997. To market, to market: the changing face of grocery retailing. *Long Range Planning,* 30**,** 836-846.

KENHOVE, P. V. & DE WULF, K. 2000. Income and time pressure: a person-situation grocery retail typology. *The International Review of Retail, Distribution and Consumer Research,* 10**,** 149-166.

KIM, S. & PARK, H. 2013. Effects of various characteristics of social commerce (s-commerce) on consumers' trust and trust performance. *International Journal of Information Management,* 33**,** 318-332.

KITCHIN, R. 2013. Big data and human geography: Opportunities, challenges and risks. *Dialogues in Human Geography,* 3**,** 262-267.

KOHIJOKI, A.-M. 2011. The effect of aging on consumer disadvantage in grocery retail services among the Finnish elderly. *Journal of Retailing and Consumer Services,* 18**,** 370-377.

KRIDER, R. E. & PUTLER, D. S. 2013. Which birds of a feather flock together? Clustering and avoidance patterns of similar retail outlets. *Geographical Analysis,* 45**,** 123-149.

LANGSTON, P., CLARKE, G. P. & CLARKE, D. B. 1997. Retail saturation, retail location, and retail competition: an analysis of British grocery retailing. *Environment and Planning A,* 29**,** 77-104.

LANGSTON, P., CLARKE, G. P. & CLARKE, D. B. 1998. Retail saturation: the debate in the mid-1990s. *Environment and Planning A,* 30**,** 49-66.

LARSON, N. I., STORY, M., EISENBERG, M. E. & NEUMARK-SZTAINER, D. 2006. Food preparation and purchasing roles among adolescents: associations with sociodemographic characteristics and diet quality. *Journal of the American Dietetic Association,* 106**,** 211-218.

LESZCZYC, P. T. P., SINHA, A. & SAHGAL, A. 2004. The effect of multi-purpose shopping on pricing and location strategy for grocery stores. *Journal of Retailing,* 80**,** 85-99.

LEW, A. & MCKERCHER, B. 2006. Modeling tourist movements: A local destination analysis. *Annals of tourism research,* 33**,** 403-423.

LINGRAS, P., HOGO, M., SNOREK, M. & WEST, C. 2005. Temporal analysis of clusters of supermarket customers: conventional versus interval set approach. *Information Sciences,* 172**,** 215-240.

LITTLE, J., ILBERY, B. & WATTS, D. 2009. Gender, consumption and the relocalisation of food: A research agenda. *Sociologia Ruralis,* 49**,** 201-217.

LOVELACE, R., BIRKIN, M., CROSS, P. & CLARKE, M. 2016. From big noise to big data: Toward the verification of large data sets for understanding regional retail flows. *Geographical Analysis,* 48**,** 59-81.

LOWE, M. & WRIGLEY, N. 1996. Chapter 1 - Towards the new retail geography. *In:* WRIGLEY, N. & LOWE, M. (eds.) *Retailing, consumption and capital: towards the new retail geography.* Harlow: Longman.

MAITY, M. & DASS, M. 2014. Consumer decision-making across modern and traditional channels: E-commerce, m-commerce, in-store. *Decision Support Systems,* 61**,** 34-46.

MALLESON, N. & BIRKIN, M. 2013a. Generating Individual Behavioural Routines from Massive Social Data for the Simulation of Urban Dynamics. *Proceedings of the European Conference on Complex Systems* 849-855.

MALLESON, N. & BIRKIN, M. 2013b. Estimating Individual Behaviour from Massive Social Data for An Urban Agent-Based Model. *GeoSimulation: Modeling Social Phenomena in Spatial Context***,** 23-29.

MALLESON, N. & BIRKIN, M. New Insights into Individual Activity Spaces using Crowd-Sourced Big Data. Proceedings of IEEE Big Data Science, 2014 Berkeley.

MARSDEN, T. & WRIGLEY, N. 1996. Chapter 2 - Retailing, the food system and the regulatory state. *In:* WRIGLEY, N. & LOWE, M. (eds.) *Retailing, consumption and capital: towards the new retail geography.* Harlow: Longman.

MARTIN, D. Conference proceedings: Mapping population over time: New ways of exploiting open data. Open Data Masterclass, 2011a University of Southampton. Economic and Social Research Council

MARTIN, D. 2011b. Mapping UK population over time. *Open Data Workshop.* University of Southapton: Economic and Social Research Council

MARTIN, D., COCKINGS, S. & HARFOOT, A. 2013. Development of a geographical framework for census workplace data. *Journal of the Royal Statistical Society: Series A (Statistics in Society),* 176**,** 585-602.

MARTIN, D., COCKINGS, S. & LEUNG, S. 2015. Developing a flexible framework for spatiotemporal population modeling. *Annals of the Association of American Geographers,* 105**,** 754-772.

MAURI, C. 2003. Card loyalty. A new emerging issue in grocery retailing. *Journal of Retailing and Consumer Services*, 10**,** 13-25.

MCPHERSON, T. N., RUSH, J. F., KHALSA, H., IVEY, A. & BROWN, M. J. 2006. A day-night population exchange model for better exposure and consequence management assessments. *Proceedings of the 86th AMS annual meeting, sixth symposium on the urban environment, Atlanta, GA.*

MINT. 2014. Retail sector gains big from Big Data. *Mint* p.Newspaper Article.

MINTEL. 2010. *ALCOHOL MANUFACTURERS DRINK IN PROFITS FROM AT-HOME CONSUMPTION, REPORTS MINTEL* [Online]. http://www.mintel.com/press-centre/food-and-drink/alcohol-manufacturers-drink-in-profits-from-at-home-consumption-reports-mintel - blank. [Accessed 7.6. 2015].

MINTEL 2012. UK Ethnic Food Market. London: Mintel Group Ltd.

MITCHELL, B. 2014. Workplace Zones: A new geography for workplace statistics. *In:* STATISTICS, O. F. N. (ed.). Titchfield: ONS.

MORTIMER, C., BARRATT, C. & CODY, C. 2015. Annual Halifax Pocket Money Survey. *In:* TEAM, H. P. (ed.) *Children's pocket money falls for the second tear in a row.* lloyds banking group: Halifax

MULLIGAN, G. F. 1983. Consumer demand and multipurpose shopping behavior. *Geographical Analysis,* 15**,** 76-81.

MULLIGAN, G. F. 1987. Consumer travel behavior: Extensions of a multipurpose shopping model. *Geographical Analysis,* 19**,** 364-375.

MURAD, A. A. 2003. Creating a GIS application for retail centers in Jeddah city. *International Journal of Applied Earth Observation and Geoinformation,* 4**,** 329-338.

NAKAYA, T., FOTHERINGHAM, A. S., HANAOKA, K., CLARKE, G., BALLAS, D. & YANO, K. 2007. Combining microsimulation and spatial interaction models for retail location analysis. *Journal of Geographical Systems,* 9**,** 345-369.

NESS, M., GORTON, M. & KUZNESOF, S. 2002. The student food shopper: Segmentation on the basis of attitudes to store features and shopping behaviour. *British Food Journal,* 104**,** 506-525.

NEWING, A., CLARKE, G. & CLARKE, M. 2013a. Identifying seasonal variations in store-level visitor grocery demand. *International Journal of Retail & Distribution Management,* 41**,** 477-492.

NEWING, A., CLARKE, G. & CLARKE, M. 2013b. Visitor expenditure estimation for grocery store location planning: a case study of Cornwall. *The International Review of Retail, Distribution and Consumer Research,* 23**,** 221-244.

NEWING, A., CLARKE, G. & CLARKE, M. 2014a. Exploring small area demand for grocery retailers in tourist areas. *Tourism Economics,* 20**,** 407-427.

NEWING, A., CLARKE, G. P. & CLARKE, M. 2014b. Developing and Applying a Disaggregated Retail Location Model with Extended Retail Demand Estimations. *Geographical Analysis,* 47**,** 219-239.

NOBLE, S. M., GRIFFITH, D. A. & WEINBERGER, M. G. 2005. Consumer derived utilitarian value and channel utilization in a multi-channel retail context. *Journal of Business Research,* 58**,** 1643-1651.

O'KELLY, M. E. 1981. A model of the demand for retail facilities, incorporating multistop, multipurpose trips. *Geographical Analysis,* 13**,** 134-148.

O'KELLY, M. E. 1983. Multipurpose Shopping Trips and the Size of Retail Facilities1. *Annals of the Association of American Geographers,* 73**,** 231-239.

O'SULLIVAN, D. & PERRY, G. L. W. 2013. *Spatial simulation: exploring pattern and process,* Chichester, West Sussex, John Wiley & Sons, Ltd.

OH, J. Y.-J., CHENG, C.-K., LEHTO, X. Y. & O'LEARY, J. T. 2004. Predictors of tourists' shopping behaviour: Examination of socio-demographic characteristics and trip typologies. *Journal of Vacation Marketing,* 10**,** 308-319.

ONS. 2011a. *2011 Area classifications (OAC)* [Online]. Newport: Office for National Statistics Available: http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/ns-2011-area-classifications/index.html [Accessed].

ONS 2011b. 2011 Census. Newport: Office for National Statistics.

ONS 2014a. 2011 Census Glossary of Terms. *2011: User guide.* Office of National Statistics.

ONS 2014b. 2011 Census variable and classification information: Part 2. *Part 2 - Dataset/table information.* Office of National Statistics.

ONS 2014c. 2012 - based Subnational Population Projections. Newport: Office for National Statistics.

ONS 2014d. Population Estimates for the UK 1951-2013. Newport: Office for National Statistics.

ONS 2015a. Family Spending: A Report on the 2015 Living Cost and Food Survey (LCF) - Table A51. *In:* HORSFIELD, G. (ed.). Newport: Office for National Statistics.

ONS 2015b. Long-Term International Migration: Ending 2014. *Migration Statistics Quarterly Report.* Newport: Office for National Statistics.

OPENSHAW, S. 1977. Optimal zoning systems for spatial interaction models. *Environment and Planning A,* 9**,** 169-184.

ORRELL, S., GREENWOOD, C. & LEWIS, R. 2015. GB Day Visits 2015 Survey. TNS: Visit England, Visit Scotland, Visit Wales.

OTTO, J. R. & CHUNG, Q. 2000. A framework for cyber-enhanced retailing: Integrating e-commerce retailing with brick-and-mortar retailing. *Electronic Markets,* 10**,** 185-191.

PACIONE, M. 2009. *Urban geography: a global perspective,* London, Routledge.

PALLANT, J. 2010. *SPSS survival manual: a step by step guide to data analysis using SPSS,* Maidenhead, McGraw Hill/Open University Press.

PAPADAKI, A., HONDROS, G., SCOTT, J. A. & KAPSOKEFALOU, M. 2007. Eating habits of university students living at, or away from home in Greece. *Appetite,* 49**,** 169-176.

PATIL, T. R. & SHEREKAR, S. 2013. Performance analysis of Naive Bayes and J48 classification algorithm for data classification. *International Journal of Computer Science and Applications,* 6**,** 256-261.

PATUELLI, R. & ARBIA, G. 2013. Editorial: Advances in the statistical modelling of spatial interaction data. *Journal of Geographical Systems,* 15**,** 229-231.

PELLETIER, J. E. & LASKA, M. N. 2013. Campus food and beverage purchases are associated with indicators of diet quality in college students living off campus. *American Journal of Health Promotion,* 28**,** 80-87.

PENTON-MEDIA 2001. Supermarket News: NEW DIRECTIONS. *Supermarket News.* New York: Penton Media, Inc., Penton Business Media, Inc.

PIACENTINI, M., HIBBERT, S. & AL-DAJANI, H. 2001. Diversity in deprivation: exploring the grocery shopping behaviour of disadvantaged consumers. *The International Review of Retail, Distribution and Consumer Research,* 11**,** 141-158.

PIPER JAFFYAY & CO. 2016. Taking stock with teens. Available: http://www.piperjaffray.com/3col.aspx?id=4035 [Accessed 11/11/2016].

PIPER, W. S. & CAPELLA, L. M. 1993. MALE GROCERY SHOPPERS'ATTITUDES AND DEMOGRAPHICS. *International Journal of Retail & Distribution Management,* 21.

POWELL, L. M., AULD, M. C., CHALOUPKA, F. J., O'MALLEY, P. M. & JOHNSTON, L. D. 2007. Associations between access to food stores and adolescent body mass index. *American Journal of Preventive Medicine,* 33**,** S301-S307.

PR NEWSWIRE. 2014. Big Data in Retail, Frost & Sullivan: Big Data Analytics Central to Customer Acquisition and Retention Strategies in Retail. *PR Newswire* p.Newspaper Article.

PRIOR, G., PHILLIPS, R. & O'DRISCOLL, C. 2014. The 2014 Food and You Survey: UK Bulletin. *In:* FSA (ed.). Food Standards Agency: TNS BMRB.

RASOULI, S. & TIMMERMANS, H. 2013a. Assessment of model uncertainty in destinations and travel forecasts of models of complex spatial shopping behaviour. *Journal of Retailing and Consumer Services,* 20**,** 139-146.

RASOULI, S. & TIMMERMANS, H. 2013b. Probabilistic forecasting of time-dependent origin–destination matrices by a complex activity-based model system: effects of model uncertainty. *International Journal of Urban Sciences,* 17**,** 350-361.

REILLY, W. J. 1931. *The law of retail gravitation,* New York, Knickerbocker Press.

REYNOLDS, J. & WOOD, S. 2010. Location decision making in retail firms: evolution and challenge. *International Journal of Retail & Distribution Management,* 38**,** 828-845.

ROBERTS, L. 2017. Convenience store numbers driven by symbol an multiple operators. IGD: The Institute of Grocery Distribution.

ROHM, A. J. & SWAMINATHAN, V. 2004. A typology of online shoppers based on shopping motivations. *Journal of Business Research,* 57**,** 748-757.

ROTHWELL, G. 2017. *Why and how retailers are winning in the food-to-go in Ireland* [Online]. IGD.com: IGD. Available: https://www.igd.com/research/retail/article/t/why-and-how-retailers-are-winning-in-food-to-go-in-ireland/i/16861 [Accessed 13/7/ 2017].

ROWELL, J. 2013. Omni-Channel Retailing. *Romanian Distribution Committee Magazine.*

ROY, J. 2004. *Spatial interaction modelling: a regional science context,* New York; Berlin, Springer.

ROY, J. R. & THILL, J.-C. 2004. Spatial interaction modelling. *Papers in Regional Science,* 83**,** 339-361.

RUGG, J., RHODES, D. & JONES, A. 2002. Studying a niche market: UK students and the private rented sector. *Housing studies,* 17**,** 289-303.

SAGIROGLU, S. & SINANC, D. 2013. Big data: A review. *International Conference on Collaboration Technologies and Systems (CTS)* 42-47.

SAVE THE STUDENT. 2016. The National Student Money Survey 2016. Available: http://www.savethestudent.org/money/student-money-survey-2016.html - earning [Accessed 11/11/2016].

SCHWANEN, T. 2004. The determinants of shopping duration on workdays in the Netherlands. *Journal of Transport Geography,* 12**,** 35-48.

SCOTTISH GOVERNMENT 2007. The relationship between off-sales and problem drinking in Scotland. *In:* RESEARCH, S. E. S. (ed.). Human Factors Analysts Limited.

SENIOR, M. L. 1979. From gravity modelling to entropy maximizing: a pedagogic guide. *Progress in Human Geography,* 3**,** 175-210.

SHEPHERD, I. D. & THOMAS, C. J. (eds.) 1980. *Retail geography: Chapter 1 - Urban Consumer Behaviour,* London: Croom Helm.

SHEPPARD, E. & PLUMMER, P. 2009. Economic Geography, Quantitative. *In:* THRIFT, R. K. (ed.) *International Encyclopedia of Human Geography.* Oxford: Elsevier.

SHOVAL, N. & ISAACSON, M. 2009. *Tourist mobility and advanced tracking technologies,* Routledge.

SMITH, C. W. 2014. Households and Household Composition in England and Wales: 2001-11. *In:* STATISTICS, O. O. N. (ed.). ONS.

SMITHERS, R. 2015. Britons shop as if they're on holiday - little and often. *The Guardian,* 21.10.2015.

SOLOMON, M. R. 2013. *Consumer behavior: buying, having, and being,* Boston, Pearson.

SOLOMON, M. R., BAMOSSY, G. J., ASKEGAARD, S. & HOGG, M. K. 2013. *Consumer behaviour: a European perspective,* Harlow, Pearson Education.

SONG, H. & LI, G. 2008. Tourism demand modelling and forecasting—A review of recent research. *Tourism Management,* 29**,** 203-220.

SYED RAKIB-UDDIN & LONGLEY, P. 2014. Temporal profile of daily sales in retail stores in London. *GISRUK.* University of Leeds.

TESCO. 2016. *Make your holiday last longer* [Online]. Tesco.com: Tesco. Available: http://www.tesco.com/holiday/ [Accessed 6th of October 2016].

THILL, J. C. & THOMAS, I. 1987. Toward Conceptualizing Trip ‐ Chaining Behavior: A Review. *Geographical Analysis,* 19**,** 1-17.

THOMPSON, C., CLARKE, G., CLARKE, M. & STILLWELL, J. 2012. Modelling the future opportunities for deep discount food retailing in the UK. *The International Review of Retail, Distribution and Consumer Research,* 22**,** 143-170.

THOMPSON, C. P. 2013. *Retail Spending and Store Location during a Recession: An Analysis of Changing Consumer Behaviour and Interaction Patterns.* Doctor of Philosophy, The University of Leeds.

TIMMERMANS, H., ARENTZE, T. & JOH, C.-H. 2002. Analysing space-time behaviour: new approaches to old problems. *Progress in Human Geography,* 26**,** 175-190.

TIMMERMANS, H., VAN DER WAERDEN, P., ALVES, M., POLAK, J., ELLIS, S., HARVEY, A. S., KUROSE, S. & ZANDEE, R. 2003. Spatial context and the complexity of daily travel patterns: an international comparison. *Journal of Transport Geography,* 11**,** 37-46.

TIMOTHY, D. J. 2005. *Shopping tourism, retailing, and leisure,* Clevedon, UK;Buffalo, NY;, Channel View Publications.

TOMLINSON, J., BULLOCK, N., DICKENS, P., STEADMAN, P. & TAYLOR, E. 1973. A model of students' daily activity patterns. *Environment and Planning,* 5**,** 231-266.

TURBAN, E., WHITESIDE, J., KING, D. & OUTLAND, J. 2017. Mobile commerce and the internet of things. *Introduction to Electronic Commerce and Social Commerce.* Springer.

UNIVERSITY OF LEEDS. 2017. *Term Dates* [Online]. University of Leeds Website: University of Leeds. Available: http://www.leeds.ac.uk/homepage/17/academic_year [Accessed 16/3/ 2017].

VICKERS, D. & REES, P. 2006. Introducing the area classification of output areas. *Population Trends,* 125**,** 15.

VISIT ENGLAND 2015a. Most visited free attractions - Yorkshire and Humberside 2015. *Annual Survey of Visits to Visitor Attractions.*

https://www.visitbritain.org/annual-survey-visits-visitor-attractions: Visit England.

VISIT ENGLAND 2015b. Most visited paid attractions - Yorkshire and Humberside 2015. *Annual Survey of Visits to Visitor Attractions.* https://www.visitbritain.org/annual-survey-visits-visitor-attractions: Visit England.

WADDINGTON, T. B. P., CLARKE, G. P., CLARKE, M. & NEWING, A. 2017. Open all hours: Spatiotemporal fluctuations in UK grocery stores and catchment area demand. *International Journal of Retailing, Distribution and Consumer Research*, 1-26.

WAGNER, G., SCHRAMM-KLEIN, H. & STEINMANN, S. 2013. Effects of cross-channel synergies and complementarity in a multichannel e-commerce system–an investigation of the interrelation of e-commerce, m-commerce and IETV-commerce. *The International Review of Retail, Distribution and Consumer Research,* 23**,** 571-581.

WANG, L. & LO, L. 2007. Immigrant grocery-shopping behavior: ethnic identity versus accessibility. *Environment and Planning A,* 39**,** 684-699.

WESTBROOK, R. A. & BLACK, W. C. 1985. A motivation-based shopper typology. *Journal of Retailing,* 61**,** 78-103.

WILSON, A. G. 1971. A family of spatial interaction models, and associated developments. *Environment and Planning,* 3**,** 1-32.

WILSON, A. G. 2000. *Complex spatial systems: the modelling foundations of urban and regional analysis,* Harlow, Pearson Education.

WILTON, D. 2006. *Long term trends and cycles in Canadian tourism*, World Tourism Organization (WTO).

WOOD, S. & BROWNE, S. 2007. Convenience store location planning and forecasting-a practical research agenda. *International Journal of Retail & Distribution Management,* 35**,** 233-255.

WOOD, S., LOWE, M. & WRIGLEY, N. 2006. Life after PPG6—recent UK food retailer responses to planning regulation tightening. *International Review of Retail, Distribution and Consumer Research,* 16**,** 23-41.

WOOD, S. & MCCARTHY, D. 2014. The UK food retail 'race for space'and market saturation: A contemporary review. *The international review of retail, distribution and consumer research,* 24**,** 121-144.

WOOD, S. & REYNOLDS, J. 2011a. The intra-firm context of retail expansion planning. *Environment and Planning A,* 43**,** 2468-2491.

WOOD, S. & REYNOLDS, J. 2011b. Managing communities and managing knowledge: strategic decision making and store network investment within retail multinationals. *Journal of Economic Geography,* 12**,** 539-565.

WOOD, S. & REYNOLDS, J. 2013. Knowledge management, organisational learning and memory in UK retail network planning. *The Service Industries Journal,* 33**,** 150-170.

WRIGLEY, N. 1987. The concentration of capital in UK grocery retailing. *Environment and Planning A,* 19**,** 1283-1288.

WRIGLEY, N. 1994. After the store wars: Towards a new area of competition in UK food retailing? *Journal of Retailing and Consumer Services,* 1**,** 5-20.

WRIGLEY, N. 1996. Chapter 6 - Sunk costs and corporate restructuring: Bristish food retailing and the property crisis. *In:* WRIGLEY, N. & LOWE, M. (eds.) *Retailing, consumption and capital: towards the new retail geography.* Harlow: Longman.

WRIGLEY, N. 2009. Retail Geographies. *In:* THRIFT, R. K. (ed.) *International Encyclopedia of Human Geography.* Oxford: Elsevier.

WRIGLEY, N., WARM, D., MARGETTS, B. & WHELAN, A. 2002. Assessing the impact of improved retail access on diet in a'food desert': a preliminary report. *Urban Studies,* 39**,** 2061-2082.

ZENITHOPTIMEDIA 2016. Student Living Index 2016. http://personal.natwest.com/personal/life-moments/student-living-index.html: Natwest.

ZHOU, L., ZHANG, P. & ZIMMERMANN, H.-D. 2013. Social commerce research: An integrated view. *Electronic commerce research and applications,* 12**,** 61-68.

12.0 Appendix

Table 12.1 Statistical analysis of the most appropriate number of cluster groups to use

	Size of cluster	10%	20%	30%	40%	50%	60%	70%	80%	90%
2 cluster		F=27.98 R^2=33.8%	F=55.04 R^2=48%	F=43.47 R^2=41%	F=27.86 R^2=35%	F=18.27 R^2=25%	F=18.76 R^2=14%	F=11.63 R^2=16%	F=11.53 R^2=19%	F=8.02 R^2=9%
1	3	11.389 A	13.33 A	14.944 A	16.556 A	17.667 A	18.5 A	19.33 A	20.222 A	21.444 A
2	44	9.659 B	11.284 B	12.587 B	13.909 B	15.152 B	16.26 B	17.30 B	18.356 B	19.78 B
3 cluster		F=20.44 R^2=42%	F=32.37 R^2=52%	F=68.68 R^2=69%	F=118.07 R^2=82%	F=158.85 R^2=86%	F=161.93 R^2=86%	F=131.05 R^2=83%	F=96 R^2=77%	F=57 R^2=77%
1	3	11.389 A	13.333 A	14.944 A	16.556 A	17.667 A	18.5 A	19.333 A	20.222 A	21.444 A
2	15	9.967 B	11.078 B	11.978 C	12.889 C	13.889 C	14.944 C	16.033 C	17.222 C	18.644 C
3	29	9.5 C	11.391 B	12.902 B	14.437 B	15.80 B	16.94 B	17.76 B	18.943 B	20.368 B
4 Cluster		F=15.37 R^2=41%	F=45.32 R^2=72%	F=95 R^2=85%	F=81 R^2=83%	F=110 R^2=87%	F=154	F=141 R^2=89%	F=114 R^2=86%	F=72 R^2=80%
1	13	9.885 B	11.872 B	13.372 B	14.962 B	16.262 B	17.32 B	18.256 B	19.154 B	20.551 B
2	14	10.036 B	11.119 C	12.012 D	12.905 D	13.869 D	14.88 D	15.941 D	17.119 D	18.5 C
3	18	9.287 C	11.065 C	12.546 C	14.056 C	15.454 C	16.667 C	17.759 C	18.806 C	20.269 B
4	2	11.5 A	13.667 A	15.417 A	16.75 A	17.83 A	18.75 A	19.58 A	20.583 A	21.833 A
5 Cluster		F=16.09 R^2=57%	F=44.4 R^2=79%	F=44.4 R^2=79%	F=75.2 R^2=87%	F=95.63 R^2=89%	F=122.1 R^2=91%	F=108.62 R^2=90%	F=86 R^2=88%	F=53 R^2=82%
1	1	11.17 AB	12.67AB	14.0 B	16.17 A	17.33 B	18.0 AB	18.83 AB	19.5 AB	20.67AB
2	14	10.036 BC	11.119 C	12.019 D	12.905 D	13.869 D	14.881 D	15.94 D	17.119 C	18.5 C
3	17	9.265 D	11.029 C	12.52 C	14.029 C	15.441 C	16.66 C	17.75 C	18.79 B	20.265 B
4	2	11.5 A	13.667 A	15.417 A	16.75 A	17.83 A	18.75 A	19.58 A	20.58 A	21.833 A
5	13	9.77 C	11.795 B	13.295 B	14.833 B	16.154 B	17.23 B	18.179 B	19.115 B	20.526 B

Notes:

1. These are decimal hours (not hours and minutes)
2. The higher the F value, the better the discrimination. Therefore although the 2-cluster solution always produced significant differences between the 2 clusters, the F value is never large and is never as large as the F values for the solutions with larger numbers of clusters.
3. The 5-cluster solution never has 5 separate letters for the post-hoc test and sometimes has two letters for some clusters, hence the results are not clear cut.

Table 12.2 Complete breakdown of the comparison between predicted cluster group and identified cluster membership using observed data for each test store (store IDs > 4000 are classified by The partner's as convenience store)**.**

Store ID	Predicted cluster group	Decision tree derived cluster group	Store ID	Predicted cluster group	Decision tree derived cluster group
830	1	2	4688	3	3
2108	1	2	4689	3	3
4206	1	1	4691	3	3
4242	1	1	4692	3	3
4253	1	3	4693	3	1
4699	1	1	4696	3	3
4700	1	1	4697	3	1
4707	1	3	4703	3	1
4711	1	1	4704	3	3
4747	1	3	4706	3	1
4759	1	1	4709	3	1
4761	1	1	4710	3	3
4782	1	3	4712	3	3
4793	1	1	4714	3	3
4796	1	3	4718	3	3
4887	1	3	4722	3	3
4930	1	1	4724	3	3
711	2	2	4726	3	3
713	2	2	4728	3	3
732	2	2	4729	3	3
810	2	2	4731	3	1
825	2	2	4737	3	3
832	2	2	4740	3	3
892	2	2	4745	3	3
2072	2	2	4749	3	3
2077	2	2	4750	3	3
2080	2	2	4755	3	3
2089	2	2	4762	3	3
2093	2	2	4769	3	1
2133	2	2	4773	3	3
2140	2	2	4775	3	3
2184	2	2	4781	3	3
2235	2	2	4783	3	3
2256	2	2	4791	3	3
2272	2	2	4795	3	3
4732	2	3	4797	3	3
4739	2	3	4800	3	3
4792	2	3	4468	4	4
4441	3	3	4765	4	3
4442	3	1	4816	4	4
4475	3	3	4957	4	4