DESIGNING A LOCATION MODEL FOR FACE TO FACE AND ON-LINE RETAILING FOR THE UK GROCERY MARKET

Elena Kirby-Hawkins

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds
School of Geography

October, 2016

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I am extremely grateful and thankful to a number of people without whom this work would not be completed. Firstly, I would like to express my gratitude to my supervisors Prof Graham Clarke and Prof Mark Birkin for their invaluable guidance and support during all these years. They have always responded to my needs and provided professional advice and motivation. I am very honoured to be supervised by such genuine individuals and extremely talented professionals. I am very thankful for their patience, punctilious proof reading of chapters and papers and encouragement during the presentations at the international conferences. I am also very grateful to my examiners Prof Alexis Comber (internal examiner) and Prof Alex Singleton for their constructive feedback and encouragement during the viva examination.

I feel very privileged to have been a member of School of Geography community and CSAP in particular and working with such bright and friendly individuals who always had time to provide advice and give an encouragement during tea and coffee breaks. I am particularly grateful to Andy Newing, Helen Durham, Rachel Oldroyd, Nick Lomax and Nick Hood for their advice , ideas and solutions to the problems I experienced during the SIM modelling process.

I also acknowledge the support of Economic and Social Research Council (ESRC) via the Retail Industry Business Engagement Network (RIBEN) in funding this research project.

I have thoroughly enjoyed sharing the office in old and new School of Geography buildings with my office mates who became my friends. Thank you to all my office colleagues, especially Suad Al Maji, Ferdinando Coelho de Souza, Michelle Kalamandeen, Andri Mardiah, Yunxia Wang, Nabati Ray, Rathakrishnan Kamla for your enthusiasm, very delicious cooking and support.

My special thanks are to Stephen Clarke, Holly Shulman and Nick Hood who became my friends and were always there to provide support and very valuable advice. My last thanks are to all my friends and family outside the academia who provided me with endless support and encouragement.

Abstract

The vast and rapid expansion of Internet usage has generated widespread online sales, making the UK one of the leading countries for e-commerce. Until now there has been no clear understanding or analysis of the spatial variations of online activities. Many studies have, however, examined the variance in online buying among different demographic groups usually based on limited survey information. These variations have often been explained by reference to two theories – efficiency theory and diffusion of innovations theory (Rogers, 1995). This lack of research to date is also manifest in the lack of consideration of online sales in traditional store location methodologies. The aim of this research is to establish a new model for site location which includes e-grocery shopping on the UK retail sector. Having explored the literature around the geography of e-commerce and the surveys of geodemographic usage, the thesis explores data unique to the academic sector-client's store revenue (for both physical and online channels) and customer data based on their loyalty card (interaction data). The analysis of these data sets established four major trends in the relationship between online share and store provision with insights into the substitution of online and physical channels in areas with limited accessibility to physical grocery stores. Using this information, a new, revised SIM is built and calibrated to include estimates of revenue for both face to face and online stores. It is hoped this will provide an important addition to the existing kitbag of techniques available to retail store location planners.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACORN	a Classification of Residential Neighbourhoods
BCG	Boston Consulting Group
B2B	Internet Between Businesses
B2C	Business to Consumers
CACI	Consolidated Analysis Centre, Incorporated
CRR	Centre for Retail Research
CAGR	Compound Annual Growth Rate
CRM	Customer Relationship Management
ESRC	Economic and Social Research Council
EDI	Electronic Data Interchange
IGD	Institute of Grocery Distribution
LAD	Limited Assortments Discounters
NRS	National Readership Survey
NICTs	New Information and Communication Technologies
OAC	Output Area Classification
OECD	Organisation for Economic Co-operation and Development
ONS	Office of National Statistics
RIBEN	Retail Industry Business Engagement Network
ROPO	Research Online, Purchase Offline
SIM	Spatial Interaction Model
WWW	World Wide Web

Chapter 1: Introduction - Research Aims, Objectives, Structure and Contribution of the Thesis

1.1. Background

The research undertaken in this thesis is the result of a studentship awarded by the Economic and Social Research Council (ESRC) under its Retail Industry Business Engagement Network (RIBEN). The RIBEN network aims to bring together academics and retail industry specialists to work on issues relevant to the retail industry. This research was made possible as a result of a partnership between the University of Leeds, the leading supermarket chain and CACI, an information technology company. CACI specialises in integrated marketing, location planning, network planning and technology solutions. The leading supermarket chain is a multinational grocery retailer based in the UK with the second largest share of 17% in the grocery market (IGDa, 2016).

The collaboration aims to develop spatial modelling techniques that allow retailers to identify the impact of on-line spending on physical stores (both new and existing) in order for retailers to better understand the geography of e-commerce and its impacts on traditional markets. Specifically, to date, retailers have limited knowledge of the interaction between online and in store buying and typical consumers' characteristics for both channels. This knowledge is important in order to understand the growing importance of e-commerce activity and the consequent geography of on line spending.

As will be explored in Chapter 2, e-commerce is becoming a key industry in the UK economy with consumers spending £1 online for every £5 of retail expenditure in total and the trend is growing (The Telegraph, 2015). Hitherto, the literature on e-commerce has been dominated by retail marketing studies, in order for retailers to better promote services and products to consumers. However, retailers are now beginning to understand that the e-commerce market is more sophisticated geographically – the fact that e-commerce usage is not ubiquitous across all geodemographic types means that the geography of e-commerce is fascinating, albeit complicated too! There is a clear indication that geography is significant when it comes to e-commerce activities, as spending patterns and habits depend on consumers' location, income, education and other socio and economic characteristics. Currently there is a lack of research, or development of tools such as spatial interaction models, which can predict or estimate the effects of the online activities on particular stores in a specific location.

The leading supermarket chain operates an in house Location Network and Planning team specialising in analytics in the grocery industry. Location planning teams perform a number of functions including network planning, market share mapping, logistics planning, customer profiling and competitors' analysis (Reynolds and Wood, 2010). One of the most important functions of the planning team is the evaluation of new sites and the estimation of potential revenues from the opening of new stores (Birkin et al. 2010a). Client's location planning team have robust and sophisticated modelling techniques (offering high levels of accuracy) for revenue prediction of new physical store locations with demand estimations typically driven from customer's home addresses.

With increases in customer mobility and complexities of trip-making behaviour new location models are required. For example, the rapid development of the convenience grocery market has identified the need for new modelling techniques to estimate fluctuating demand and smaller store revenues. This is the subject of another RIBEN studentship at the University of Leeds (Hood, 2016). The growth of e-commerce activity also presents new challenges to traditional modelling techniques. Client has been actively developing this channel since 1995 obtaining a 75% delivery coverage in the UK by 2014. Since March 2015 the supermarket chain has also developed a 'click and collect' service with the numbers of stores offering this service expected to double by the end of 2017. Currently, client offers 'click and collect' services at 712 grocery sites (J Sainsbury plc, 2016). The location planning team will need to identify catchment areas for the stores offering this service, to estimate potential demand and allocate resources to fulfil this service with allowances for busy collection times, i.e. Saturday mornings.

Client's in house location planning team uses a Spatial Interaction Model (SIM) to estimate revenue from new grocery sites. This technique has proved to be effective in predicting revenues for physical stores to date. Online grocery retailing presents more complex challenges for the store location team. How can they estimate demand for online retailing in particular locations. Are customers in rural locations more likely to use online channels in comparison to city dwellers? What is the profile of a typical online customer? Furthermore, how will physical stores be affected by the growth of online expenditure and, conversely, what is the impact of physical stores on online channel demand? How can they quantify the attractiveness of the online offer? This thesis attempts to help client answer these questions. In addition, it will help to add online activity into the traditional SIM methodology, something that has not been achieved to date by any researchers (as evidenced at least by literature searches). This is important as the SIM technique is one of the major location planning tools used by client's site location team. To validate the model estimated, actual sales data from 131 supermarket stores provided by the client and used in the model construction. This store data from the partner organisation will help to analyse and understand online purchase patterns and identify specific products categories that are more frequently purchased online. Moreover, the data from individual transactions and linked loyalty cards, which includes customer postcodes, will enable the identification of the spending patterns online or in stores within local catchment areas and beyond. The RIBEN studentship thus offers a unique opportunity to access and analyse data rarely seen in the academic world.

To recap, this research provides the first attempts to design a SIM location model which includes on-line grocery sales. The next section will introduce the broad aims and objectives of this thesis.

1.2. Aims and objectives

The main aim of this project is to develop new spatial modelling techniques that allow retailers to first understand the geography of online sales and then to include online sales in models of store forecasting. To achieve this aim the research intends to design a spatial interaction model which will estimate the online and face to face demand for retail products and to estimate the market share of individual retailers across both physical and online channels, i.e. to give retailers a profound understanding of their target market to enable them to develop their online and physical store offerings in an optimal fashion.

The broad aim of the investigation can be broken down into a number of key objectives, which comprise:

1. To explore geographical variation in online grocery sales in the Yorkshire and Humberside region.

2. To understand, in far more depth, the geodemographics of online customers of one particular retailer.

3. To examine the geographical variations of online activities influenced by distance to physical stores in Yorkshire and Humberside applying the data from partner organisation.

4. To explore and model the flows of people between their homes and various retailers, using a spatial interaction model to allocate demand for online and face to face buying

5. To provide recommendations on the provision of existing physical stores in particular locations and make suggestions where additional stores could be located to fulfil demand.

In addition, the research seeks to provide a more profound understanding of the geography of e-commerce activity. This will include an examination of the growth of the e-commerce industry in the UK and its future development; a discussion of products which are likely to be purchased on line and a comparative analysis of the interaction between online and instore shopping in terms of complementarity, substitution, neutrality and modification.

The objectives outlined above refer to the overall aim and intention that this project will be able to estimate online expenditure across the UK (not just Yorkshire and

Humber discussed here) and the findings of this research will be representative and can be applied to other retailers and other regions.

1.3. Thesis structure and scope

To meet the aim and objectives outlined above, the structure of the thesis is as follows. Chapter 2, as a part of the literature review, explores the growth of the e-commerce industry in the UK and outlines its scope and impact on the UK economy. Furthermore, it identifies the composition of the e-commerce market and highlights the opportunities and threats for future growth. This chapter will conclude by identifying that location is becoming increasingly important to retailers. Finally, this chapter will identify the research gaps in the literature which this thesis hopes to fill. Chapter 3 examines the spatial distribution of Internet users and the impact of spatial variables (location and shop accessibility) on e-commerce activities. This chapter will test two hypothesises: the innovation-diffusion theory versus the efficiency theory. Chapter 4 reviews existing methods used by retailers in location based analysis, identify the strengths and weaknesses of each technique and makes a comprehensive justification for the chosen methods. The following models will be compared and contrasted – analogue, regression, spatial interaction, geographical information system, ratings, econometric and agent-based models.

Chapter 5 examines the data supplied by the partner organisations CACI and major supermarket chain. The validity and size of the data will be examined in order to provide comprehensive analysis of e-commerce activity within convenience market. Chapter 6 estimates the demand for online grocery products and present the results and discussion of analysis using CACI survey data and supermarket loyalty card data. The intended analysis will identify the consumption habits of on line customers, the distribution among supermarkets and convenience channels and estimate the demand for online channels in the study area. Chapter 7 aims to establish the spatial distribution of grocery sales in the study area. To achieve that, a traditional SIM for face to face retailing will be created for study area. Based on the findings from Chapter 5 and 7, chapter 8 will present a new model that will be able to estimate demand for online and face to face purchases within a store catchment area. The location (rural or urban), product classification, customers' socio, economic and geodemographic categorization will be used to determine the spatial distribution of e-commerce activities. A few "what if" scenarios will be created. The final chapter will conclude the thesis and evaluate the developed spatial interaction model. It will outline its limitations and make some comments of its use, and offer ideas for further development, in the future.

1.4. Thesis contribution

The objectives outlined above refer to the overall aim and intention that this project will be able to estimate online expenditure across the UK, although, initially the research will concentrate on the Yorkshire and Humberside region as it offers contrasting attributes – rural/urban, affluent areas/deprived neighbourhoods, etc. Since all the major retailers hold detailed store level data for their own and their competitors' network, it is anticipated that many potential end users of the location model will have access to the required supply side data for their area of interest. Consequently, the major contribution that this project can make to the academic literature, and its major benefit in commercial applications, will be the contribution made to the understanding of online demand, which is currently under-researched, followed by the production of new types of location model for retail site forecasting. The outcomes of this research have already been presented at many international conferences with great interest shown by academics and retail industry specialists alike.

Chapter 2: Development of e-commerce and omni channel retailing in the UK

2.1. Introduction

This chapter explores the growth in e-commerce with an emphasis of the effect of social changes on online activities within the UK retail sector. First, the chapter introduces the UK e-commerce economy sector, its position within the global e-commerce industry, current trends and reflects on future developments. Section 2.2 provides definitions of the primary research subject of e-commerce, introduces the e-commerce industry in the UK and identifies possible motives behind the growth. Section 2.3 outlines the development of the e-commerce industry worldwide and provides an overview of its position within the UK industry and in the retail sector in particular. Section 2.4 discusses in more detail changing consumer attitudes and potential reasons for online channel growth. Sections 2.5 and 2.6 provide comparative analysis of e-commerce and face to face retailing and overview of the current online consumer segmentation studies. The components of the multi-channel retailing phenomenon will be examined in section 2.5. Section 2.7 outlines some key advantages of e-commerce from the consumer's perspective. Furthermore, problems created by e-commerce will be analysed in section 2.8. The concluding section (2.9) provides an overview of the UK grocery sector (section 2.9.1), including the development and growth of the online channel (Section 2.9.2) and presents an overview of the current online grocery market within multi-channel retailing.

2.2. E-commerce: definitions

The primary subject of this research is online shopping or e-commerce, which is described as a commercial activity performed on the Internet between businesses (B2B) or businesses and consumers (B2C) (Mokhtarian, 2004). In this research the term e-commerce will refer to B2C transactions. The e-shopping process includes buying and searching for goods and services online (ibid). Similarly, the phrase e-tailing will be used from time to time. This refers to electronic retailing and selling products on the Internet (OECD, 2011). The organisation for Economic Co-operation and Development (OECD) defines e-commerce as:

‘the sale or purchase of goods or services, conducted over computer networks by methods specifically designed for the purpose of receiving or placing the orders’ and ‘the goods or services are ordered by these methods, but the payment and the ultimate delivery of the goods or services do not have to be conducted online’ (OECD, 2011, p.72).

Total e-commerce sales comprise of sales made over the websites and Electronic Data Interchange (EDI) which is a direct computer to computer data transfer. In addition, researchers at the Centre for Retail Research (CRR) define e-commerce as retail sales made over the Internet (including use of mobile phones and tablets) to the final consumer with the exclusion of fuel, cooked food, holidays, tickets, insurance and gambling (CRR, 2016)

Despite the rapid and widespread Internet usage since the invention of the World Wide Web (WWW) the growth in online retail sales in the 1990s and the early 2000s was very slow due to concerns over Internet security and difficulties in site navigation (Williams, 2009). Moreover, during the initial stages the Internet was viewed as a tool for information exchange and gathering rather than a commercial medium (ibid).

There are three stages in the development of any e-commerce market. First, the so-called ‘immaturity stage’ where online market share is below 6.5% and online spending is geographically and demographically widely dispersed, with less than ten online purchases per annum per person (CRR, 2016). Italy, Poland and Spain are examples of countries which fit this development phase. However, they are expected to overcome the difficulties with Internet coverage and increase the base of regular e-shoppers very rapidly. The second stage is often referred to as ‘mid-range’ (The Netherlands, Sweden and France for example) with online market share rising from 6.5% to 9.5%, and 45% of the population making purchases more frequently via computers and mobile devices. The market reaches a ‘mature stage’ when over half of the population are regular online shoppers, making over twelve online purchases annually with market share above 9.5%. The UK, Germany and US have become mature markets in the last few years and are actually expected to slow down with further growth being generated by existing online shoppers buying more frequently and buying more expensive products.

2.3. Growth of e-commerce

The vast and rapid expansion of Internet usage has generated widespread online sales, with the UK one of the leading countries for e-commerce today. In 2015 around 7% of all retail trade worldwide was undertaken online and it is expected to double by 2019 with the UK as the European leader generating 15% of online sales compared to total retail expenditure (EMarketer, 2016). E-commerce is one of the fastest growing industries in Europe and USA achieving a growth rate of 18.4% in 2015 in Europe with almost half of the population being an online shopper on some occasion (CRR, 2016). In comparison, all other types of retailing achieved growth rates of only 1.5% to 3.5%. Thus far, the USA is the worldwide leader in online retailing with 62.3% of its population defined as an online shopper with an average annual expenditure of £1119.79 in 2015 compared to £820.05 spent by European e-shopper (ibid).

The UK is the leader in the European market with total online retail expenditure reaching £52.25billion with an annual growth rate of 16.2% in 2015 (CRR, 2016). Figure 2.1 demonstrates the online shares of retail trade for eight European countries for the period 2014-2016 with the three leading countries, UK, France and Germany achieving 81.5% of all online retail sales in these eight countries. Despite the overall growth of online retail sales, the European mean has remained relatively low, rising from 7.2% in 2014 to 9.4% in 2016, although researchers at the Centre for Retail Research predict that online retail sales will achieve an 18% market share in each of these eight European Countries in the near future (CRR, 2016).

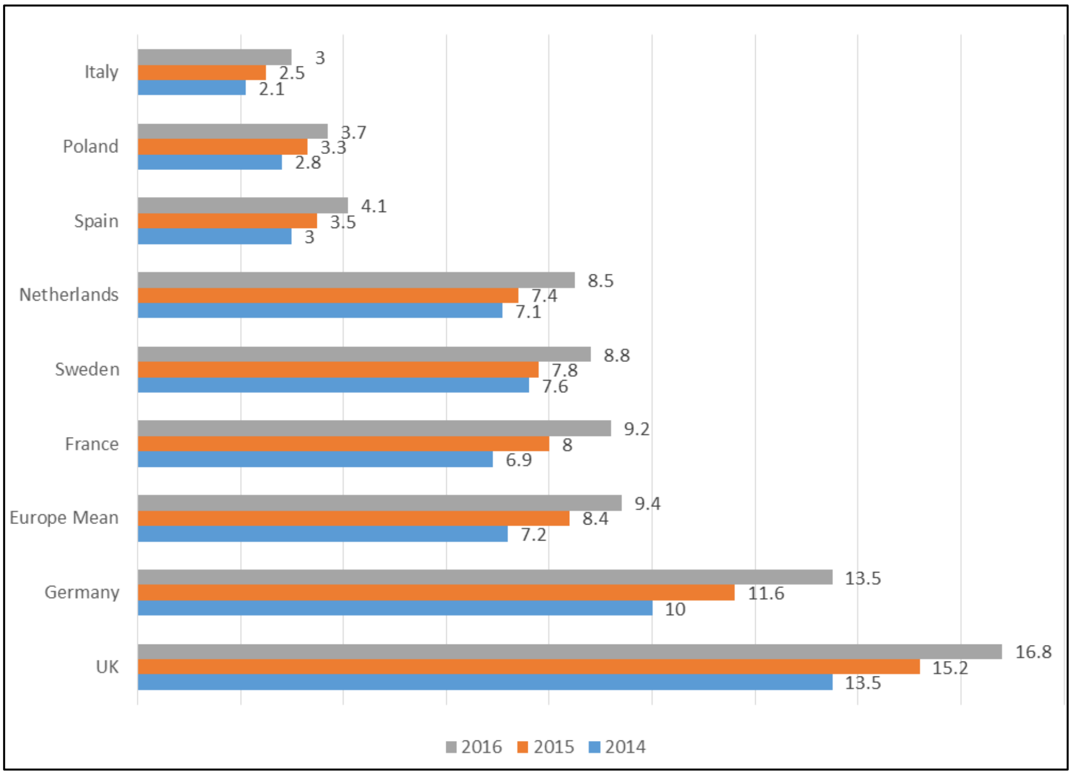


Figure 2.1. Online share of Retail Trade 2014-2016: Source: Centre for Retail Research, 2016

In terms of national and global GDPs share, the e-commerce sector is relatively low with values ranging from 1.18% on a global scale to 1.52% in the US in 2015 (Figure 2.2). The growth of the e-commerce sector as a share of GDP has been relatively stable with an annual growth rate of approximately 0.1% worldwide. In the last few years, the e-commerce sector has expanded its share of GDP by approximately 0.2% and in 2018 will reach 1.61% share of global GDP, which is an increase of 1.07% for the nine year period (CRR, 2016).

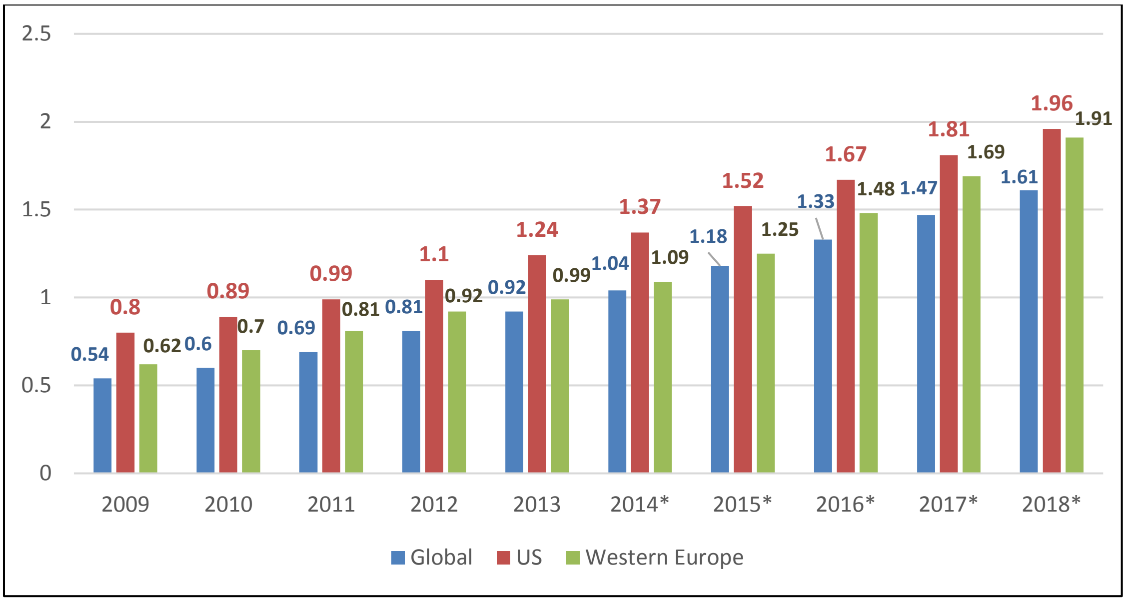


Figure 2.2. E-commerce as percentage of GDP from 2009-2018: Source: Statista, 2016

In 2015 the UK e-commerce sector accounted for 6.1% of total GDP with a Compound Annual Growth Rate (CAGR) of 10.9% for the period 2010-2016 (Watershed Publishing, 2016; Consultancy.UK, 2016).

An interesting question is why does the UK lead the e-commerce sector across Europe? The answer lies in the combination of high Internet penetration (87%), a well-developed e-commerce infrastructure, a very competitive market and high consumer confidence in the security of using credit cards for on-line payments. (Watershed Publishing, 2016; Webinterpret 2015). In 2015, UK consumers spent £114 billion online which is an increase of 11% compared to the previous year (Retail Gazette, 2016). UK consumers thus spend £1 online in every £5 of retail expenditure and the trend is growing (The Telegraph, 2015).

In 2014 the e-commerce sector (website sales) contributed £199 billion to UK business turnover which is an increase of 80% since 2009 and outperformed every other major economic sector (ONS, 2015a). Figure 2.3 demonstrates the breakdown of online sales by different UK business sectors in 2014.

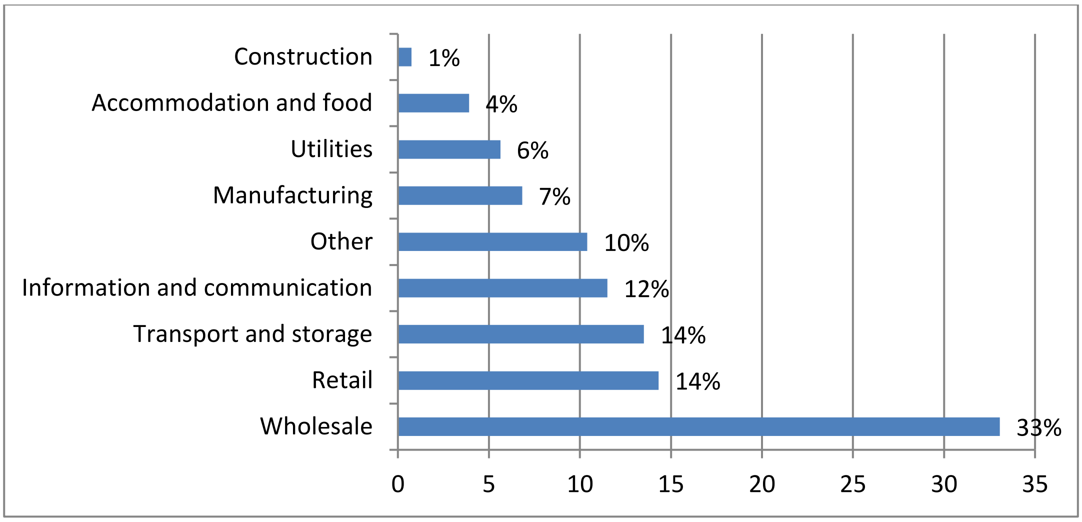


Figure 2.3. UK e-commerce sales via a website, by industry sector: Source: ONS, 2015a

The construction industry has the lowest value of goods sold online with only 1% compared to the highest percentage of 33% in the wholesale sector. Retail and transport industries contributed 14% each to total website expenditure. Accommodation, food, utilities and manufacturing sectors have low usage of online channels, with less than 10% each, although these sectors of the economy have doubled their online sales since 2009 (ONS, 2015b). Within the retail industry the highest percentage of Internet sales in 2014 naturally came from within the 'non-store' retailing sector (which includes mail order, catalogues and stores trading mostly over the Internet) with 69% of all sales being completed online in this category (Figure 2.4). The lowest percentage for online retail sales is within the food sector accounting for only 3.7% of the total grocery expenditure in 2014, but is now estimated to be 6% according to Mintel (2016)

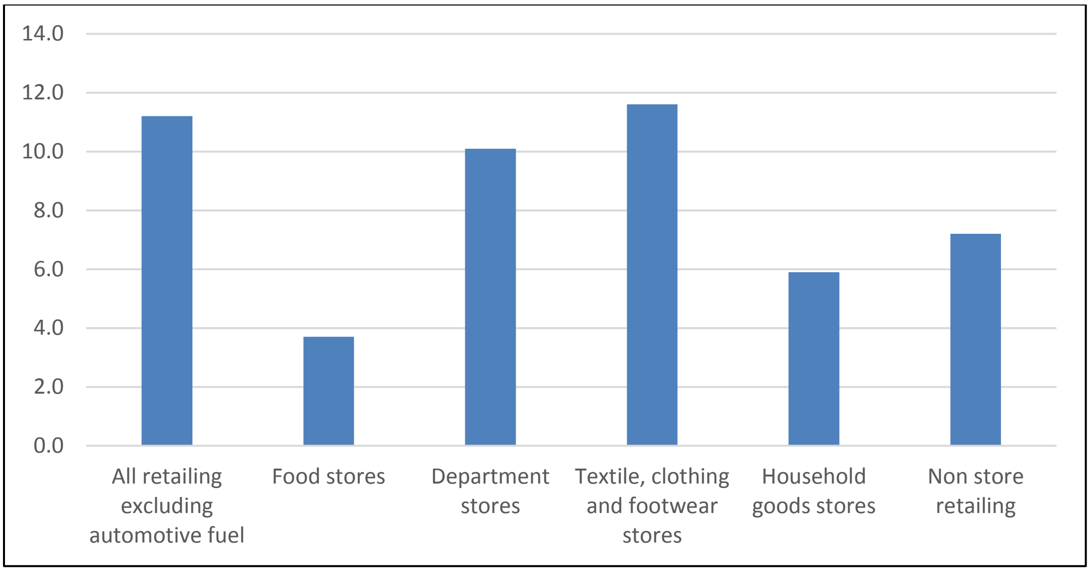


Figure 2.4. Percentage of GB Internet sales in each retail sector 2014: Source: ONS, 2014

Table 2.1 demonstrates the dynamics of online sales within the UK retail industry for the period 2011-2014. The highest growth of 3.8% and 3.5% in online sales during this period was within the textile, clothing and footwear and department stores sectors, compared to 'other' stores selling jewellery, toys, sporting goods and books in the non-food category where online sales decreased by 0.4%. In general, online sales increased by3% in the total retail market between 2011 and 2014, with the highest increase in non-store retailing of almost 8%.

Table 2.1. Annual proportion of total sales made online (%)

Category	2011	2012	2013	2014
All retailing	8.3	9.3	10.4	11.2
All food	**2.7**	**3.1**	**3.4**	**3.7**
All non-food	**6.9**	**7.7**	**8.4**	**8.8**
Department stores	6.6	7.7	9.5	10.1
Textile, clothing and footwear stores	7.8	9.2	10.2	11.6
Household goods stores	5.1	5.9	5.7	5.9
Other stores	7.6	7.5	7.9	7.2
Non-store retailing	**62.1**	**66.4**	**67.9**	**69.4**

Source: ONS, 2015b

Thus, the rapid growth and importance of e-commerce in the UK industry is evident. The future of e-commerce will be examined in the next section.

The rapid expansion of e-commerce is unquestionable, which instigates further important questions – at what rate will e-commerce continue to grow and has it reached its maximum level? There are a number of different predictions. According to leading consultants IMRG Capgemini, by 2020 the total spent online will have doubled compared with the 2010 figures (cited in Palmer 2010). At this time growth will have reached its maximum and is then predicted to decelerate to only 6% growth per year. Analysts from the Boston Consulting Group expect e-commerce activities in the UK to reach £221billion by 2016 (a growth rate of 11% per annum) which will outperform the existing major players in this market such as the USA and China (Palmer, 2012). The UK will remain an important world player in e-commerce, with its share in total retail sales increasing from 14.5% in 2015 to 19.3% by 2019 (EMarketer, 2016)

2.4. Reasons for online growth

The growth of e-commerce sales is closely related to advances in ICT, changing social trends and on-going economic conditions, although the most important factor is often said to be proficient technological provision, e.g. efficient network coverage, high Internet speed and sophisticated devices. The OECD[1] (2008) established that there has been a direct link between broadband expansion and increase in e-commerce activities. The exponential growth of broadband subscription between 2000 and 2006 (from 0 to 25 per cent) had an impact on e-commerce activities in the UK (Economics, 2015). The UK government recognised the importance of high speed broadband national coverage with the plan to provide superfast broadband (with the speed of 24Mbps or more) for 95% of the UK premises by December 2017 (Gov.UK, 2015). Although, the spatial disparities of broadband connection and access still exist with 18% of UK population never used the Internet in 2013 (Riddlesden and Singleton, 2014). The 'digital divide' or 'digital differentiation' phenomenon described in the works of Paul Longley (2008) and Hargittai (2008) changed from Internet access to more complex social differences due to the quality of the broadband. Riddlesden and Singleton (2014) in their work, established that rural areas receive slower broadband

[1]Organisation for Economic Co-operation and Development

connection with urban areas suffering from 'bottlenecking' of data traffic during pick times of the day. A more detailed analysis of broadband usage in the UK will be described in Chapter 3.

Another important growth factor has been the increased use of e-commerce via tablets and smartphones, often labelled m-commerce (shopping via mobile device). According to Ofcom (2012), UK consumers spend more money on line buying via mobile phone Internet usage than any other nation in the world (Thomas, 2012). British consumers spend over £1000 a year on products purchased via the mobile phone; the majority of them are ITunes, cinema tickets and clothes (Thomas, 2012). James Thickett, Ofcom's Director of Research, observes that British people are known to be the quickest to accept new innovations and they are the highest mobile phone shoppers due to easy access via devices like smartphones (ibid). In 2015, m-commerce increased by 42% compared to the previous year, with 24% of all online sales being made via smartphone devices and the trend is growing (Retail Gazette, 2016). Figure 2.5 shows the predicted growth of e-commerce sales by 8.3%to 2019, which will be largely driven by growth through m-commerce, with the share reaching over 19% of total retail expenditure by 2019. The share of m-commerce sales in total e-commerce sales will reach almost 44% in 2019 which is an increase of 12% compared to 2014. Moreover, in 2015 there was an increase in m-commerce usage among all age groups (see Chapter 3 for more discussion on geodemographics).

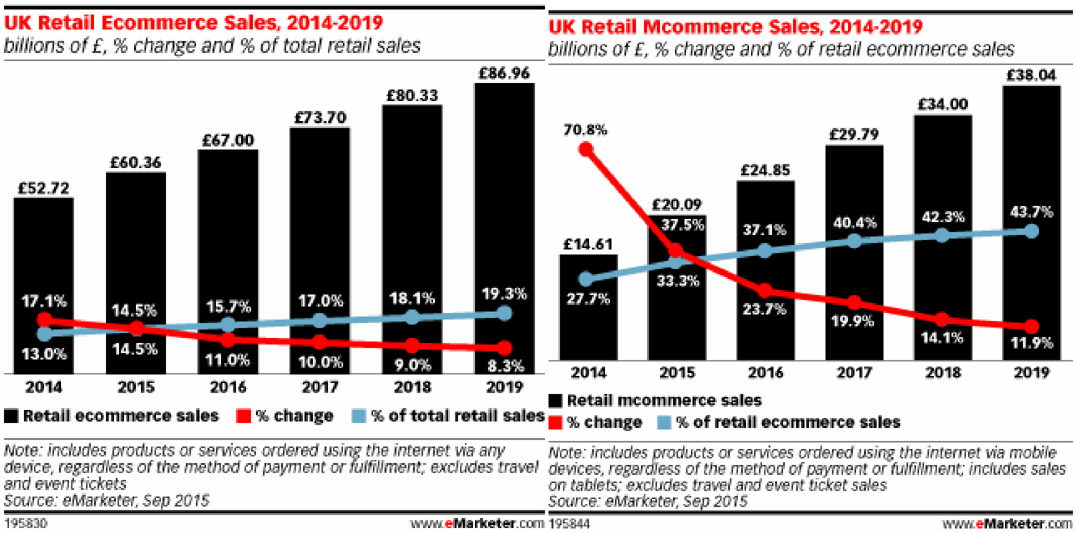


Figure 2.5. UK Ecommerce sales and Mcommerce sales forecast 2014-2019: Source: EMarketer,2016

Moreover, modern society is becoming increasingly socially connected and we are now witnessing the ever increasing use of social media platforms which create new opportunities for businesses to reach their target market and influence the way customers make purchasing decisions. Many scholars have noted that social media has an advantage over traditional marketing strategies and it's more effective in influencing customers (Singh et al, 2012; Heinonen, 2011; Mangold and Faulds, 2009). In 2015 5% of all e-commerce sales were attributed to social media with Facebook as being a leader, generating 85% of all social media e-commerce orders (Tradeglobal, 2015). According to the same source social media has a huge potential for growth in online sales, with 90% of consumers saying they trust their friends products' recommendations compared to only 33% who trust advertisements.

2.5. E-commerce and face to face retailing

According to Verdict, in association with SAS, a new retail model is emerging where retailers need to provide a combination of large stores (with full ranges), smaller satellite stores with limited or sample stock in more convenient locations, plus a full online range of e-commerce facilities, including 'click and collect'(Verdict, 2013). Such multi-channel modes of operation are increasingly commonplace. Debenhams, for example, is already operating such a model with its department style large stores, its smaller Desire units, 'click and collect', m-commerce and kiosk facilities. Unquestionably, other retailers will follow this example. The necessity of a multi- or omni-channel model demonstrates that retailers must be flexible and adapt to emerging customer needs. Omni-channel means 'every channel' (Sealey, 2013). The main difference between omni and multi-channel retailing is that latter refers to a retailer's presence on a majority of channels, whereas omni-channel is concerned with making all channels work effortlessly together. This new concept provides customers with multiple choices of channels to find information, purchase and receive goods in the most convenient way for the customer. Figure 2.6 demonstrates the interaction between customers and stores within the omni-channel system.

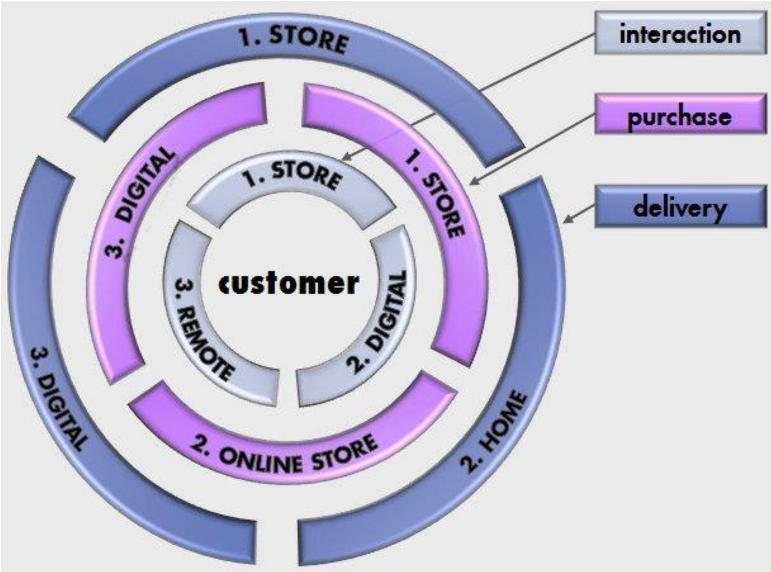


Figure 2.6. Omni-channel components: Source: Ispira, 2012

This new development relates to an innovative approach in creating a customer experience rather than satisfying a pure functional retail transaction. In order to provide this experience it is essential for businesses to create a unified brand experience across all channels. To transfer from multi to omni-channel systems, retailers need to establish the following key elements. First, companies must have an effective digital platform. Currently, 50% of the UK companies' websites do not comply with EU cookie law, which is designed to protect online users privacy and obligates websites to receive consent from visitors to store information (The Guardian, 2015). Secondly, an omni-channel system requires a new organisational structure within all departments from marketing to operations working in close collaboration. Finally, retailers must know their customers in order to develop an appropriate channel for a specific target market.

Omni channel retailing closely relates to the shopping process and it is useful to understand more about the components of the shopping process and how e-shopping integrates with it. The growth of e-commerce creates new challenges for existing consumer decision making models of the future. "Social apponomics" is a new term in e-commerce, suggesting that businesses need to create new business models with regards to social media, bespoke applications and customer's ever changing needs (Anderson et al., 2010). To be successful, businesses need to build online customer life-time value. Booz and Co (2010) identified six elements of this new 'value' (see Figure 2.7).

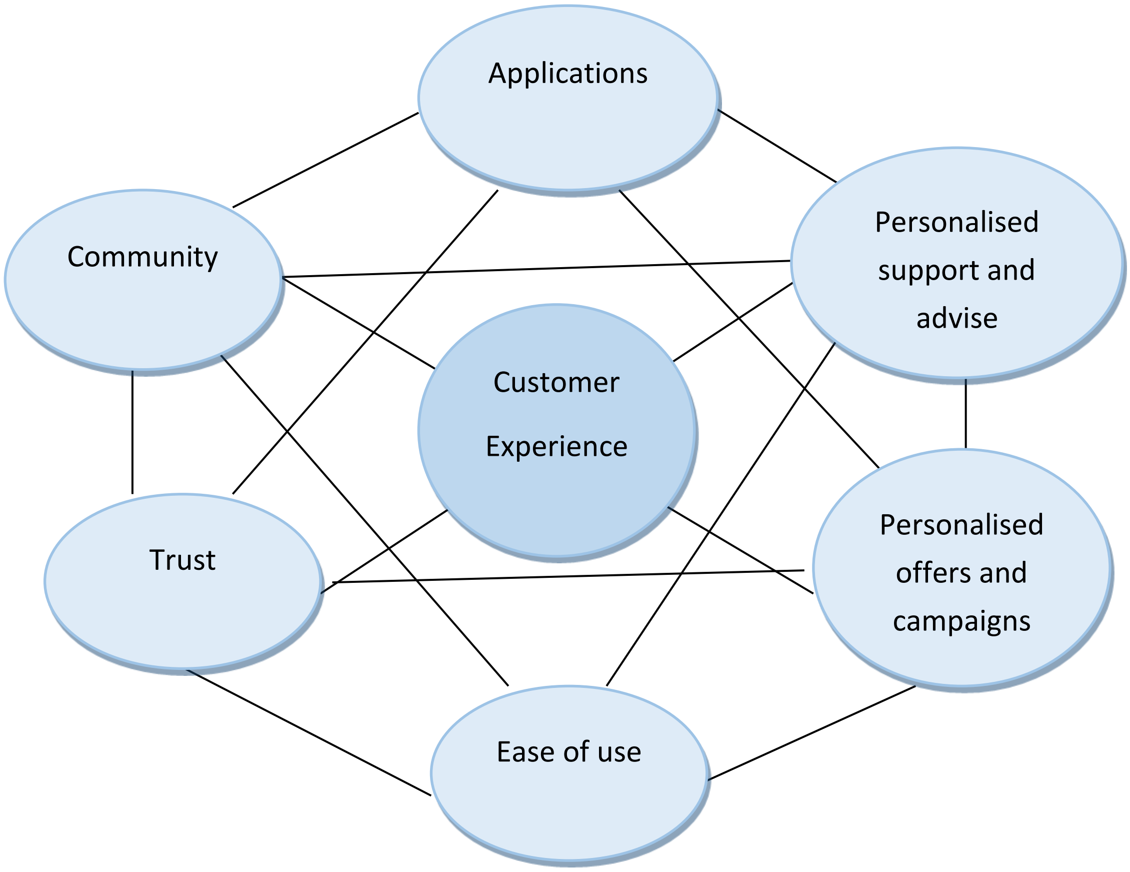


Figure 2.7. Six elements for creating Online Customer Lifetime Value: Source : Booz and Co (Anderson et al, 2010)

The next generation of business models should merge traditional elements of customer experience, i.e. trust, ease of use, personalised offer together with these new social apponomics characteristics, i.e. community, bespoke advice and applications. The supermarkets, airlines and other retailers created loyalty programmes many years ago but the appearance of applications stores and social network websites has changed the loyalty concept (Anderson, et al., 2010). The successful sites (with a deep understanding of the new model) embrace emotional, technical and social needs of the 21st century generation customers. Applications include software tools allowing users to manipulate their own data. Bespoke advice is provided by experts but more frequently (and cost free) by other users as well. Bespoke offers and campaigns are often supported by customer relationship management (CRM) to ensure that all customers receive personalised messages and advertisements, which generates more revenues for businesses. To satisfy customers' ever increasing needs modern sites must have a quick, user friendly and attractive visual appearance. Trust refers to offering an efficient returns policy and being transparent in business policies and offers. The last (but not the least) feature is the community, which allows website users to exchange their views and ideas with the knowledge that they will be recognised by other users with the same outlook, which helps to create a sense of belonging. Anderson et al (2010) conclude in their paper, that online businesses will need to be ever more location specific and target specific customer groups. Netflix, the largest online rental service, is an example of an online business which incorporates all six elements of customer lifetime values. All businesses must follow suit in order to survive and succeed in a new environment. "Companies that succeed in monetizing online access will be those that develop customers for life online" (Anderson et al. 2010, p.12).

The growth of e-commerce, as hinted above, has naturally had a major impact on shopping behaviour and changed the way in which some consumers behave, especially in relation to the mix of on-line versus face-to-face shopping. Retail analysts have identified four different ways in which on-line buying and visiting a physical store interact: substitution, complementarily, modification and neutrality (Mokhtarian 2004, Weltevreden 2006, Farag 2006). It's useful to look at each of these in turn:

Substitution

The substitution phenomenon refers to when on-line purchases completely replace a trip to a physical store. Many research findings confirm this trend. For example, Dixon and Marston (2002) identified that 28% (amid the sample of 450 UK consumers in a town in southeast England) had replaced many in-store purchases.

Complementarity

Weltevreden (2006) segregates the complementarity effect into two categories – enhancement and efficiency. The former occurs when e-activities directly induce physical purchases, e.g. promotions and advertising. The latter refers to e-shopping enhancement, e.g. the provision of click and collect services allows consumers to save on delivery charges and eliminates safety issues of payment on line by allowing wary customers to order goods on line and complete the purchase in store. He notes that these two categories are difficult to separate in practice and further research is required to explore the efficiency effect, as current studies apply an enhancement effect when referring to this complementarity phenomenon (see more discussion on this below).

Modification

E-commerce can change the behaviour attitude of shoppers. For example, a consumer may search for product information on-line but then travel to a particular store as a result of this search. In addition, modification refers to a change of mode of transport, duration and destination (Weltevreden, 2006). It suggests that the shopping process will be more efficient as buyers will be better prepared and will spend less time on purchasing the product and/or on travelling to and from a shop.

Neutrality

In this instance e-shopping does not affect in store shopping and vice versa. This phenomenon largely varies among different product categories and depends on the frequency of online purchases (Weltevreden, 2006). For example, the effect of e-commerce on the shoes and jewellery categories will be minimal as these products are purchased less frequently on line compared with face to face.

These stages form a shopping cycle and not all stages will necessarily be completed for every purchase. In addition, some elements will be repeated again until the required product is purchased (Salomon and Koppelman, 1988). Today, many consumers use a mixture of on-line and physical shopping (Farag, 2006). Couclelis (2004) describes this process as "fragmentation" of well-established activities, e.g. work or leisure and "recombination" in an innovative approach, i.e. for any or all of these shopping cycle stages shopping activity can occur in the workplace, while travelling, doing household tasks, etc. She identifies three types of consumers during the three stages of shopping process – before purchase, during and after. To assess local store viability, she further categorises shoppers with regards to their shopping modes - remote (on line) and local (in store). The types of shoppers are: "the traditional shopper (local/local/local), the cybernaut (remote/remote/remote), the good citizen (remote/local/remote) and the free rider (local/remote/local). The traditional shopper will search for goods, purchase them and use post sales services at local retailers. The free rider is completely opposite, using local traders to test the goods and return goods to them but making purchases on-line. The former kind of shopper benefits the local retailer more than the latter which might cause problems for local vendors. The price conscious consumers will frequently purchase online as it normally offers greater choice of goods at more competitive prices.

Mokhtarian (2004) agrees that e-shopping can substitute for all the stages of the shopping process as e-tailers search for new ways to reach their customers. Encouraging buyer's desires via pop-banners for example, makes it possible to test products on-line (with options such as of music sampling or the virtual fitting room). On-line retailers can also provide a convenient payments systems and offer a returns policy to encourage on-line sales.

The buyers' decision of which channel to choose will depend on the following four dimensions – individual characteristics, product characteristics, shopping mode characteristics and shopping motives. Consumers are not interested in NICTs for themselves, but how they can enhance their shopping experience and facilitate the making of better decisions (Burke, 2002). Figure 2.8 summarises the consumer decision making process for online and in store channel retailing. Each dimension offers advantages and disadvantages for both channels. The consumer's purchasing decision will be based on three key factors – time, quality and cost (Wilson-Jeanselme and Reynolds, 2005 (also see Clarke et al, 2012).

Figure 2.8. Consumer Decision Process (CDP): Source: Farag, 2006

Weltevreden (2006) states that substitution and complementarity are more likely to occur for products that are more frequently purchased on-line. He notes that generally on line browsing has a positive effect on physical shopping, while on-line buying negatively effects in-store shopping. There is no clear indication of which four types of e-commerce impact the greatest on physical shopping. According to Jupiter Communications only 6% of online sales will not be generated to the detriment of in store sales (Mintel, 2015b). In her paper Patricia Mokhtarian states neither in store or online shopping will dominate the other (Mokhtarian, 2004). She concludes that consumers will use both channels and retailers need to market both modes of retailing. According to Langston (2011), retail stores are not going out of fashion and the Internet offers many opportunities for retailers to increase their sales. Langston (2011) even suggests that the substitution process will have an opposite effect as pure e-tailers started opening physical stores since stores also drive online sales.

Interestingly, in their recent projects with two major retailers, CACI has established that their online sales increased as a result of store presence due to the effect of brand awareness and the existence of a "click and collect" services. They applied the Retail Footprint Catchment Model, which enables retailers to predict an increase of online sales with the opening of a new store. Figure 2.9 demonstrates this trend with online sales falling with the further distance from the store.

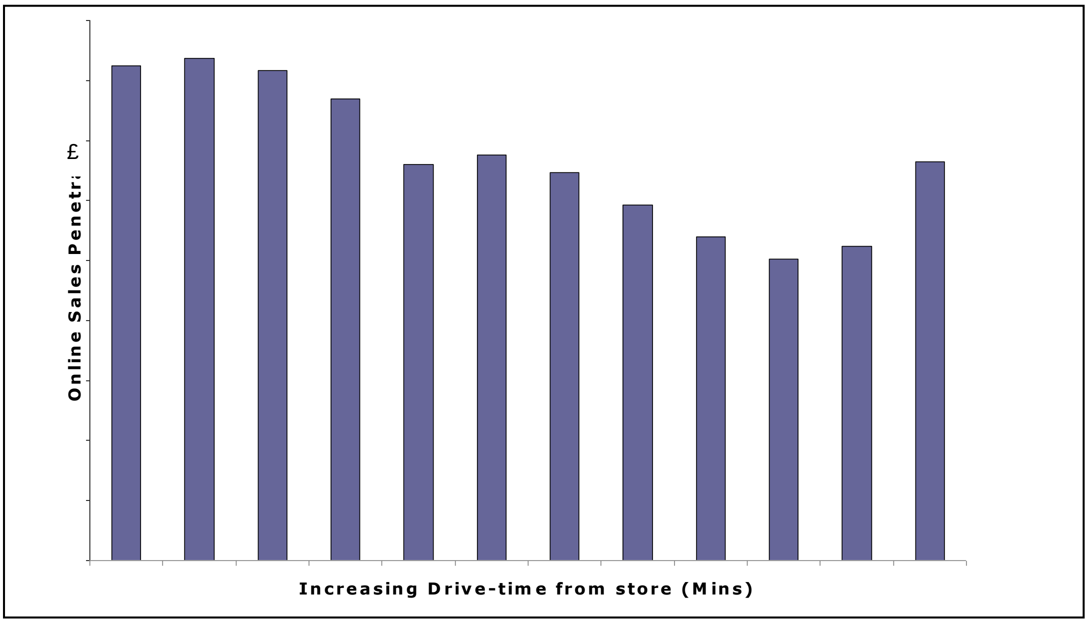


Figure 2.9. Online Sales Decreasing with Drive-time from Store. Source: Langston, 2011

Many commentators have noted that goods and products that involve little risk or effort are purchased more often online (Cao and Mokhtarian, 2005; Rotem and Salomon, 2004; Farag 2006). However, consumers can become more confident and make more high value purchases online (Datamonitor, 2012). After consumers have successfully purchased low cost products, for instance books and CDs, they become more adventurous with their online experience and venture into buying widely differentiated goods, i.e. clothing and grocery (Mintel, 2013). According to the same source, product and seller brands have a robust impact on online sales. Novice Internet users tend to buy on-line from the brands they have developed trust for in physical stores (Mintel, 2013). It is suggested that the elderly population will tend to purchase from the brands they know as they are late adopters of on-line shopping.

Cliquet (2006) relates the definition of little effort to the concept of distance, which is becoming increasingly inconsistent with a changing perception from static and linear to constantly changing variable relating to another volatile physics component – time. Braudel (1986) described this phenomenon as “the true measure of distance is the speed of human movement” (Cliquet, 2006, p.33). Another complication is perception of time between customers either due to travelling at different times of the days or to interpersonal and cultural traits. The travel modes also influence the perception of time, e.g. perception of car users will differ from customers travelling by public transport (Allemand, 2001). Leo (2000) describes the trips are no longer radial but more like loops (Dion and Cliquet, 2006)

The complexity related to purchase is represented in Figure 2.10

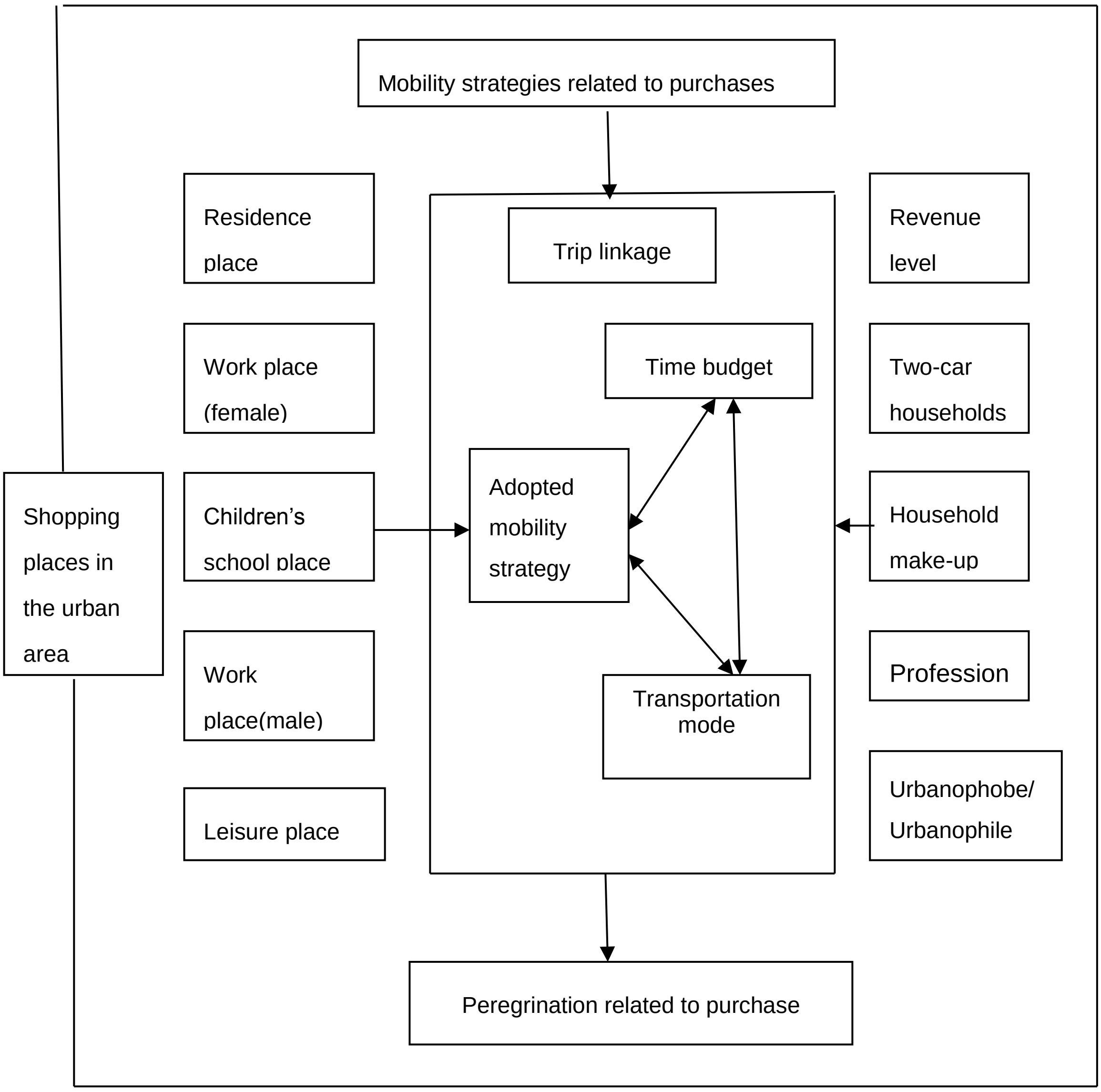


Figure 2.10. Mobility behaviour systems related to the frequenting of shopping place. Source: Dess, 2001 (Dion and Cliquet, 2006, p.34)

Dion and Cliquet (2006) use term 'peregrination' to describe the modern customer shopping trips given the chaotic nature of humans' movements. They distinguished two types of mobility – *insular* or more routine movements and *network* trips which are less concentrated in space and time. Interestingly, despite the more sophisticated and faster transportation modes (cars and buses), the average daily trip time of 55min by French consumer remained stable between 1982 and 1994. The modern consumer travels longer distances but without spending more time on shopping trips. The researchers noted that nowadays we are living in the *postfordian* model of movements where each individual moves at any possible opportunity and not at specific times compared to *fordian* model where individuals take specific trip at the specific time.

The shopping trips became more opportunistic rather than planned with customers making their shopping on the way to work, home, etc.

It is important to explore the interaction between searching for products prior to purchase as many consumers start their product information search before they go to physical stores (Ward and Morganovsky, 2002). The research is concerned with two possible variations – consumers will search on line prior travelling to the shops or they will explore physical stores before they complete a purchase on line. Weltevreden (2006) in his study based on a sample of 3218 Dutch Internet users, identified that majority of consumers (88%) browse city centre shops before making an in store purchase, with the most frequent categories electronics, domestic appliances, jewellery and telecoms. However, the Internet is the second most used channel for information gathering prior to the purchase in city centre shops (in similar product categories of electronics, domestic appliances, software and travel). The least searched categories are grocery, shoes and clothing. Not surprisingly, 92% of all e-shoppers searched on line before making an e-purchase, with the most popular products CDs, videos, DVDs, books, travel, outer clothing, computer hardware, second-hand goods and collectibles. Second-hand goods, collectibles, groceries, health and personal care and travel were the least searched products in city centres when purchased on line. Significantly, in his paper, Weltevreden (2006) discovered that e-shoppers will use different channels for information gathering, whereas an in store shopper will use on average only one channel.

The comparative analysis of offline and online channels does not show huge advantages of one channel over another. The conclusion is that these two modes of shopping are becoming increasingly intertwined with the difficulty to separate and measure the effect of the ROPO (and vice versa) phenomenon on overall retail sales. The choice of a channel will depend on types of customers or consumer segmentation.

2.6. Review of online consumer segmentation studies

Many scholars have performed consumer segmentation studies, although there are a limited number of segmentation and demographic studies which have addressed omni-channel buying behaviour. Bhatnagar and Ghose (2004) applied latent class modelling to segregate online customers based on their shopping behaviour within various product categories. Consumers were segmented based on their age, education, gender, income, marital status and Internet usage. The research found that price alone is not a vital attribute for online shopping and the majority of the respondents searching online did not complete the purchase on the websites due to the perceived security risks associated with the online purchases.

Konus et al (2008) implemented in their research the search for products stage and identified three segments of consumers based on their hedonic and economic variables – multichannel enthusiasts, uninvolved consumers and physical stores focused consumers. The research concentrated on consumers' psychographic characteristics and did not find sufficient evidence between any socio demographics and shopping behaviour. They studied six psychographic characteristics: price consciousness, shopping enjoyment, innovativeness, motivation to conform, loyalty and time pressure. Enjoyment relates to social aspects of shopping, which will not be very important for grocery shopping. Innovative customers will try to find different new products and will have more extensive searches to explore various products and options. Motivation to conform relates to consumers being assured from external sources (friends) that their intended purchase decision is correct. The more loyal customers use less channels and spend less time exploring alternatives. Switching between brands takes effort and time. Time conscious consumers will try to utilise the channel with the highest efficiency.

Lieke van Delft (2013) segmented consumers based on three product categories (fashion, personal care and grocery) and five phases of the shopping journey (stimulation, search for information, purchase, delivery, after purchase service). Two-step clustering analysis was applied based on the five respondents' demographics characteristics – age, gender, household income, employment and education. The following consumer clusters were identified within the grocery product category.

1. Omni-channel grocery shoppers. Consumers within this group sometimes search online but prefer physical stores to buy their groceries. The profile of the customer belonging to this cluster is brand and retailer loyal, 50years old or below, in full or part time job, highly educated with high incomes. Moreover,

these consumers are innovative and search for goods and services on social media websites and applications while shopping in physical stores.

2. Offline targeted grocery shoppers. These customers are more likely to be male 50years old or above, who do not use the Internet to buy groceries or search for products, although, they can be encouraged to buy or search for products in stores.
3. Offline grazers. These consumers are very similar to the previous cluster with the difference that shoppers in this group are more likely to be females 50 years old or above.

These online customer segmentations provide an insight into potential online customer locations based on their demographic characteristics. For example, gender was identified as one of the major demographic factors relating to the acceptance and usage of new technologies. The literature review indicates that Internet is a male dominated environment including online shopping (Elgar, 2008), although, other demographic characteristics are important in building demand for online expenditure as will be described in Chapter 3. The knowledge of online customer profiles is an essential part in designing a site location model which will estimate online and face to face expenditure and revenues. The next section looks at the growth of e-commerce in relation to the perceived advantages for the consumer.

2.7. Advantages of e-commerce to the consumer

What are the major advantages of online retailing to the consumer? There are four major factors:

1. **Price.** According to ECommera the deciding factor for 6 out of every 10 online consumers is price (Palmer, 2010). Pure on-line retailers can offer lower prices due to lower costs of market entry and operating costs (Mokhtarian, 2004).

2. **Delivery.** According to ECommera convenient delivery options was the deciding factor for 5 out of every 10 respondents (Palmer, 2010). Free delivery is also a key factor in purchasing online for 8 out of 10 customers (ibid).

3. Recommendations/trust/knowledge of brand. Some commentators note that consumers are not always driven by lowest prices even when buying goods such as DVDs and books. Other factors can be more important in making consumers choice – trust, brand loyalty and habits (Mokhtarian, 2004). Over 70% of UK shoppers buy online following recommendations and nearly 50% of all on line sales were made as the shoppers were already familiar with the brand (Palmer, 2010). The UK favourite high street brands, i.e. Marks and Spencer and John Lewis, scored fourth and seventh places respectively in YouGov survey among British online shoppers (Palmer, 2012). Amazon came first in that poll.

4. Convenience and speed. The online channel wins greatly over physical stores in this category as it offers unlimited 24/7 access for searching and buying goods (Mokhtarian, 2004).

In terms of a better shopping experience, the following key factors, encourage consumers to shop online more frequently (Datamonitor, 2012):

1. Free delivery
2. Site navigation
3. Free returns policy
4. Easy payment system
5. Pictures , videos and interactivity
6. Personalisation of the website

Despite these apparent advantages of online shopping, especially cost effectiveness, physical stores have the upper hand in certain areas of the shopping experience. Salomon and Koppelman (1988) describe the following seven dimensions:

1. Sensory information. Products which require physical contact prior to a buying decision will not be purchased on-line. However, they admit that when virtual reality becomes more sophisticated consumers will buy sensory goods on line.
2. Tangibility of the shopping environment. This dimension relates to shoppers loyalty, trust and habit to retailers known in the community rather than anonymous on-line retailers.
3. Immediate possession refers to instant access to products which should be balanced against the time spent travelling to shops and time waiting for

delivery. However, there is a danger that certain goods might not be in stock, e.g. correct size or colour. To gain advantage in this dimension many on-line retailers offer click and collect services.

4. Social interaction refers to the social aspect and the enjoyment of shopping.
5. Entertainment. Many analysts observe the recreational aspect of the shopping experience. The majority of modern retail centres combine entertainment, e.g. cinema, game arcades, etc. with the variety of shops.
6. Movement. Some researchers refer to a person's desire to go shopping for its own sake to satisfy the need just to get out of the house (Mokhtarian and Solomon, 2001).
7. Trip chaining, i.e. many consumers combine their shopping trip with travel for other purposes.

2.8. Impacts of e-commerce on the built environment

It is clear that e-commerce has become a principal contributor to the UK economy. However, it raises concerns that the economy becomes excessively reliant on this sector and that it has grown at the expense of other sectors. The closure of businesses, empty shops, unsightly vacancies on the high streets (with signs "to let") are frequent images around the UK. In the opinion of many commentators this trend is not entirely due to recent economic downturns but a change in consumers' habits in favour of online shopping rather than going into a physical store. Langston (2011) suggested that companies will eventually close 20% of their stores if 20% of their sales are completed via an online channel. However, some companies, i.e. John Lewis, have expanded their physical presence despite the fact that their online channel is very successful also. Mike Jervis, insolvency partner and retail specialist at PwC, agrees that Internet sales alone cannot be attributed to store closures as bankrupt retailers had an excessive number of stores within sufficient multi-channel activities (BBC News, 2013).

Despite Wrigley and Dolega (2011) suggesting that UK high street shops (with diversity and corporate-food-store entry characteristics) have been more resilient to economic downturn, as many as four in ten shops have been estimated to have closed down due to the online alternatives (Palmer, 2012). Recent causalities of this trend have been

HMV, Blockbuster, Comet and Jessops. According to the Local Data Company, in 2012 there was a total of 1779 (net) shop closures across 500 British town centres with the South East suffering the most with 376 net closures in the region (BBC News, 2013). Figure 2.11 demonstrates that the prime losers have been retailers selling computer games, health foods, cards, bookshops and music stores. According to Statista (2013), for example, the digital share of music sales in the UK rose from 0.8% in 2004 to almost 32% by 2011 (BBC News, 2013). Many analysts attribute the store closures also to the economic downturn and competition from supermarkets selling non-food products (ibid).

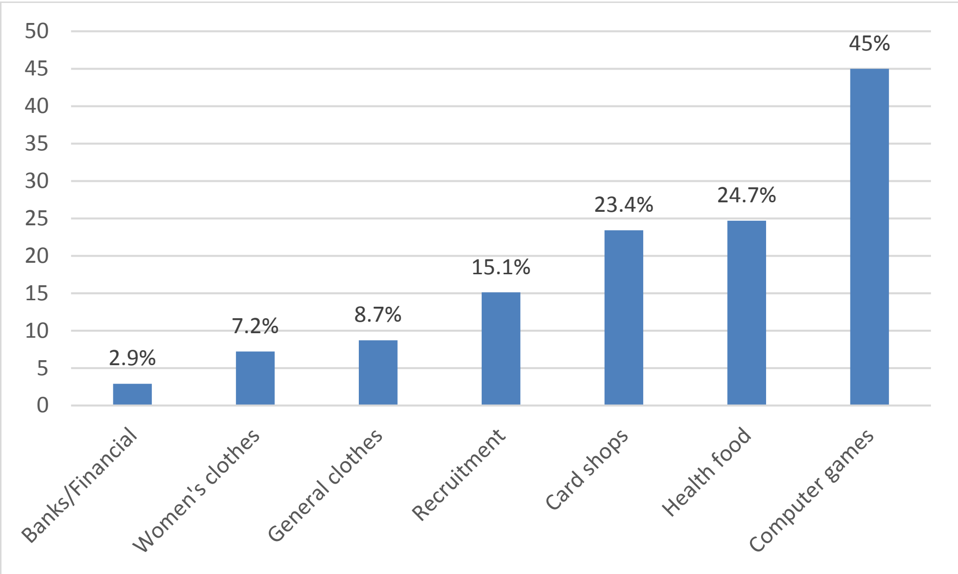


Figure 2.11. High streets store closures in 2012: Source: Local Data Company (BBC News, 2013).

Moreover, the role of the town centres is changing from retail orientated to more recreational leisure purposes with a growing number of coffee bars and restaurants (Datamonitor, 2012). In 2010 sales in town centres increased by just 0.2%, while sales of out of town retail centres increased by 1.6% over the same period. Non-store sales grew by 10.4%.

Indisputably, Internet sales have had a huge impact on the traditional brick and mortar high street retail sector. However, the effect is not inevitably negative. In the last few years a new trend is emerging in the retail sector, with pure players (online only retailers) opening physical stores to expand their presence in offline environments also. The examples include multiple fashion retailers Spar too and Bonobos, and even

Amazon opened their first physical book store in USA in November 2015 (LS Retail, 2016).

Moreover, as noted above, some traditional retailers are benefiting from the online channel as a significant research tool for customers prior to them making a purchase in a physical store. According to a Boston Consulting Group (BCG) estimation in 2010, among the top 20 leading countries for on-line sales, ROPO (research online, purchase offline) contributed almost twice as much to the total retail economy compared to online sales only (Watershed Publishing, 2016). The reverse occurrence (research offline, purchase online) has a significant impact on the overall retail experience also, with over a third of UK consumers reporting researching in store prior to completing a purchase online (Econsultancy, 2014).

The emergence of e-commerce has therefore created new challenges and opportunities for modern retailers with the need to adapt to the new retail environment with consumers becoming increasingly technologically savvy, expecting businesses to satisfy their demanding needs to use all channels to make a knowledgeable choice (from purchasing bread in a local supermarket to booking a holiday in an exotic destination). The next section will explore the trends associated with the grocery market in particular.

2.9. Development of e-commerce in the UK grocery sector

In 2014, the UK online grocery market was worth £7.7 billion or an estimated 4.4% of the total grocery market (IGD, 2016a). Despite the current low overall market share, the online grocery market has doubled in value since 2009 and it is predicted that it will be worth £16.9billion or 8.3% of the total grocery sector by 2019. Hence it is becoming the principal driver of the UK grocery market (IGD, 2016a). The consumers' increasing demand for flexibility, convenience and reliability (discussed above in relation to the UK sector as a whole) places a great pressure on the grocery retailers to create a seamless approach to the shopping experience through the availability of many different channels – in-store, online, click and collect (collectively again known as omni-channel retailing).

2.9.1 Overview of the grocery retail sector

Before dealing with e-commerce it is useful to provide a brief summary of the changing UK grocery market as a whole. The grocery retailer is defined as a merchant that primary sells food along with household and pet goods (Competition Commission, 2008). The modern supermarkets or superstores have been labelled as much more than just retailer opportunities: many scholars describe them as 'retail theatres' with 'themed environments', 'fantasy urbanism' and 'Carnivalization', where customers can shop, bank, be treated, dressed and socialise (Wrigley and Lowe 2002). Moreover, the UK grocery retail industry is an influential political and economic force, being the largest private sector employer and the largest manufacturing sector in the UK.

The grocery sector is the major contributor to the UK economy with over 50% of total retail expenditure being spent on groceries (IGD, 2016b). Analysts at the Institute of Grocery Distribution (IGD) forecast that by 2021 the UK grocery market will be worth £196.9 billion which is an increase of almost 10% compared to £179.2 billion grocery sales in 2016. However, when analysing the dynamics of the grocery sales over the last 5 years, from 2011 to 2016, the food industry can be seen to have experienced a continuous fall in annual growth rates from 4.7% in 2011 to 0.4% in 2015 (The National Farm Research Unit, 2014). The structure of the UK current grocery industry is represented in Figure 2.12.

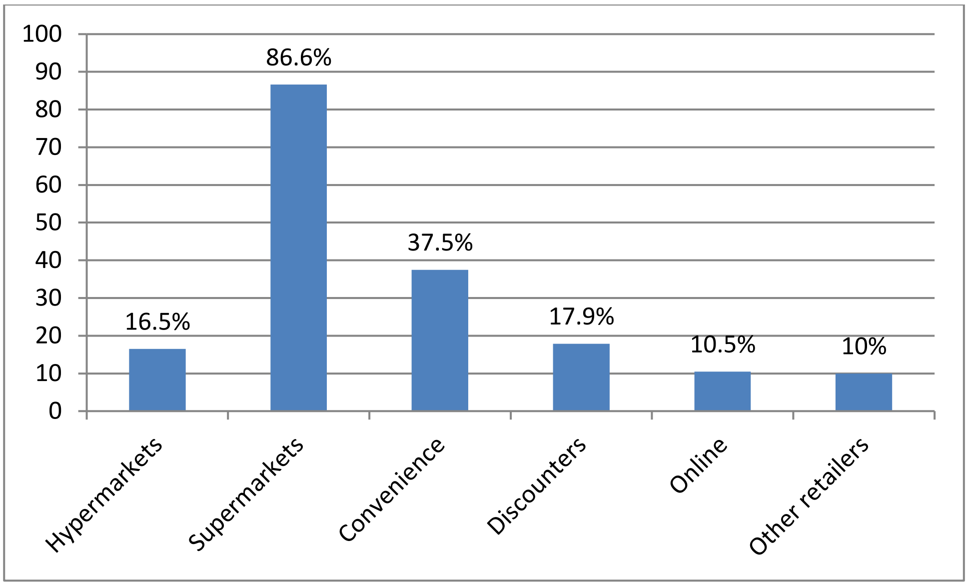


Figure 2.12. The composition of the UK grocery sector: Source: IGD, 2016

The supermarkets dominate the market with almost 50% of all groceries purchased through this channel. The convenience stores are the second most popular channel with 21% of the share. Discounters and online channels have relatively low shares of 10% and 6% respectively. The IGD provides the following categorization of the UK grocery channels:

1. Hypermarkets. The largest type of grocery stores with the size of over 60,000 sq ft and an extensive range of food and non-food lines.

2. Supermarkets. Large grocery stores with extensive food lines and a small range of non-food goods, with a size of between 3,000 and 60,000 sq ft

3. Convenience stores (c-stores). Small grocery stores with the sales area of less than 3,000 sq ft with longer opening hours offering at least seven product categories but limited household goods

4. Limited Assortments Discounters (LAD) or Discounters. High street or out of town retailers selling groceries at lower prices compared to the major food supermarket chains. Examples - German based Lidl and Aldi, Poundland and B&M.

5. Other retailers. High street small retailers with less than 3,000 sq ft sales area including newsagents, off licenses, bakeries and department stores selling food as a side line of their predominantly non-food business orientation

6. Online. Groceries ordered online for home delivery or for pick up at the local supermarket, i.e. 'click and collect' service

Figure 2.13 shows changes in the grocery sector predicted for the next five years with a total market growth of over 23% by 2020. The highest growth is expected to be within convenience, discounters and online channels with superstores losing their market share by 2% by 2020.

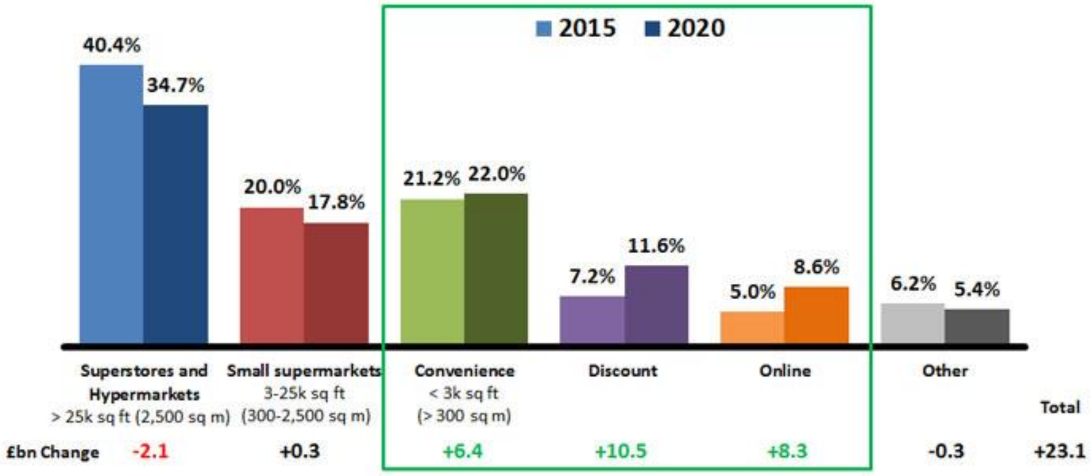


Figure 2.13. Changes in the UK grocery industry 2015-2020. Source: IGD, 2016a

The market is dominated by four leading supermarkets chain or the “big four” – Tesco, Morrisons, Asda and Sainsbury’s, collectively accounting for 70% of the market share and 60% of the total grocery floorspace (KantarWorldPanel, 2016; Hughes et al, 2009). Figure2.14 demonstrates the distribution of the grocery market among the UK’s leading grocery retailers with Tesco taking a substantial lead with 28% of market share followed by Sainsbury’s and Asda with the similar market shares of 16%. The discounters Aldi and Lidl have a combined share of 9% compared to one of the “big four” Morrisons at 11%.

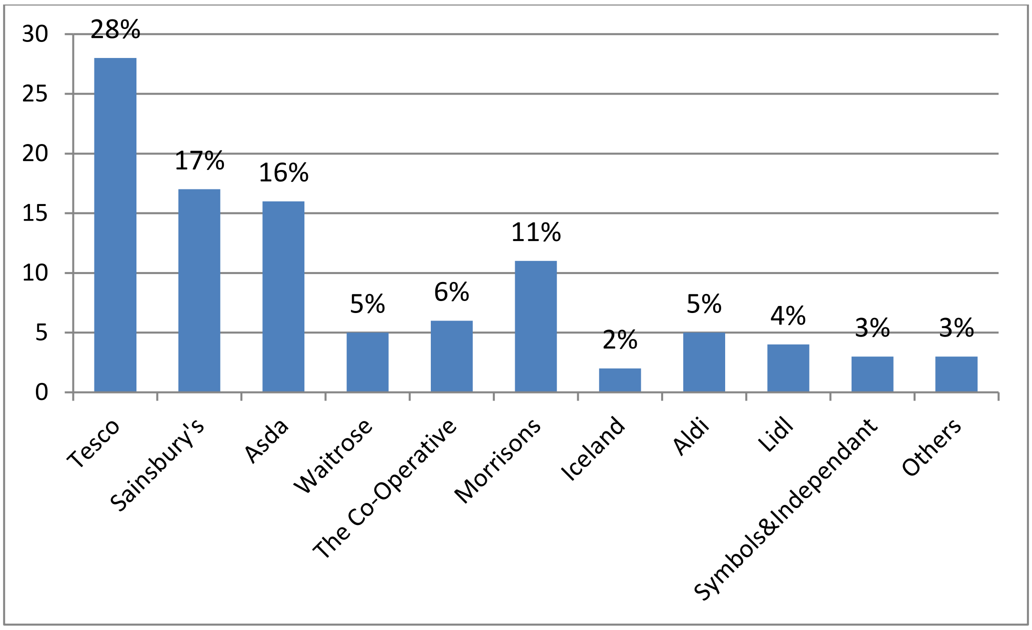


Figure 2.14. Grocery Market Share 2015 Source: Kantar Worldpanel, 2016

Tesco has twice the floorspace compared to its closest rival Sainsbury's, with an 11% advantage in market share. At the same time, Tesco's operating margins of 6% are higher than average for an industry where 3.6% to 4.5% is the norm (Competition Commission, 2008).

The current structure of the grocery retailing started to develop in the 1950s with the shift from service to self-service and the growth of smaller supermarkets. In the 1970s and 1980s large, mainly out of towns supermarkets were built, with a small numbers of major retailers obtaining more e market share and consequently, declining numbers of the smaller independent stores (Competition Commission, 2008). From 1960 to 2000 the number of large and mid-sized grocery stores more than trebled from 2000 to 6300 stores. Between 2000 and 2007 the super large format supermarkets (with over 25000 sq ft) experienced an annual growth of 3% compared with 1% growth for smaller format stores. As a result of mergers, acquisitions, organic growth and 'store wars', the top five largest food retailers (Sainsbury, Tesco, Safeway, Asda, Somerfield) controlled over half of the whole retail market share at the end of 1990s. They became very powerful politically and economically in the UK. In the 1990s their power was threatened by increased market penetration from European 'deep' discounters retailers, e.g. Netto and Aldi rapidly expanding due to the limited discounted subsector present at that time in the UK (Kwik Save being the only major incumbent), and the low prices offered in times of intensifying recession. The market leaders were forced to reconsider their strategies and introduced ‘price fighter’ brands to be competitive and loyalty card schemes to retain customer loyalty. They recognised the consumer trend towards more frequent but smaller shopping baskets and moved into convenience market in the 2000s with Sainsbury’s purchasing the chain of Jackson’s stores in 2004 and Tesco’s acquisition of T&S small stores in 2002 (Seely, 2012). They repositioned the strategies, moved with demand and continued to be the market leaders (see Hood et al 2016 for more details)

The grocery industry is changing. The recent financial crises have made customers develop new habits, e.g. being more price conscious and shopping at alternative discount stores. Consequently, the big four’s profits are falling, with Sainsbury’s declaring in 2015 a pre-tax loss of £72mn, and Tesco of £6.4bn (The Telegraph, 2015). Furthermore, the heavy investments into large format grocery stores in order to expand their presence has resulted in seemingly underused floorspace across all “big four” supermarkets with profits falling per sq ft of grocery floorspace. For example, Sainsbury’s declared in 2016 one in four of their supermarkets to having space under-utilised, which totals to 6% of the total retail space. The supermarkets have increasingly recognised the problem and are now finding alternative ways to utilise the available space by offering more non-food product ranges and letting the empty areas to other shops. For example, the purchase of Argos by Sainsbury’s in 2016 allowed the first 10 Argos stores to be opened in Sainsbury’s supermarkets in 2016.

As with the retail market as a whole, the grocery market is becoming a very competitive environment with new technologies presenting new opportunities for businesses to enter the market. The existing market leaders have experienced further pressure from the development of new entrants to online retailing, i.e. Ocado, a pure e-retail grocery offer and Amazon using their existing online platform to compete with the traditional format grocery.

2.9.2. Growth of UK e-commerce in grocery sector

The development of grocery online channels began in 1985 with Tesco and Asda offering a home shopping service through the computer, although it was not until 1995 when the online channel began its rapid growth with Tesco recognising its importance and making it the core of their business strategy (Digital Foodie, 2013). Tesco became a pioneer of the grocery online channel in the UK and initially chose a store based fulfilment model to service the online orders. The main advantages of this approach are that it utilises the existing resources (in store staff picking online orders) and minimises the investment into a new business operation for which demand is uncertain (Fernie, 2010). Moreover, it allowed the rapid geographical expansion of the online grocery market. However, the store-based model had a few drawbacks. For example, product selection and their availability were limited by the physical size of the store and subsequently the substitution rates could be as high as 10% (McClellan, 2003). Moreover, the inconvenience and logistics complications of picking the online orders in-store meant the traditional in-store service consequently suffered from a decline in standards of service. Asda and Sainsbury's applied different models, including building dedicated warehouses or dotcom stores to service on-line only orders. The warehouse approach is believed to be more effective allowing better management of stock inventory and a more extensive product range. However, this model requires a considerable initial investment of building costly warehouses in exchange for a very modest long term return of 6% in the total online market share. Thus, not surprisingly, Asda and Sainsbury's had to close many of these just a few years after opening them. Tesco waited until scale economies could facilitate a more successful warehouse style operation, eventually opening its first warehouse (or 'dark' store) in 2006.

Despite the Tesco's first mover rewards, on-line grocery retailing is now an increasingly competitive environment with new players entering the market and capturing market share.

Figure 2.15 shows the breakdown of market share for the on-line grocery market. Tesco is the principal UK market leader in both face to face and e-commerce channels:

around 28% market share in the traditional UK grocery market and almost 40% in online sales (Kantar Worldpanel, 2016; Econsultancy, 2014, Figure 2.15). Figure 2.15 also shows the predicted change from 2009-2019 in online shares among the groceries retailers.

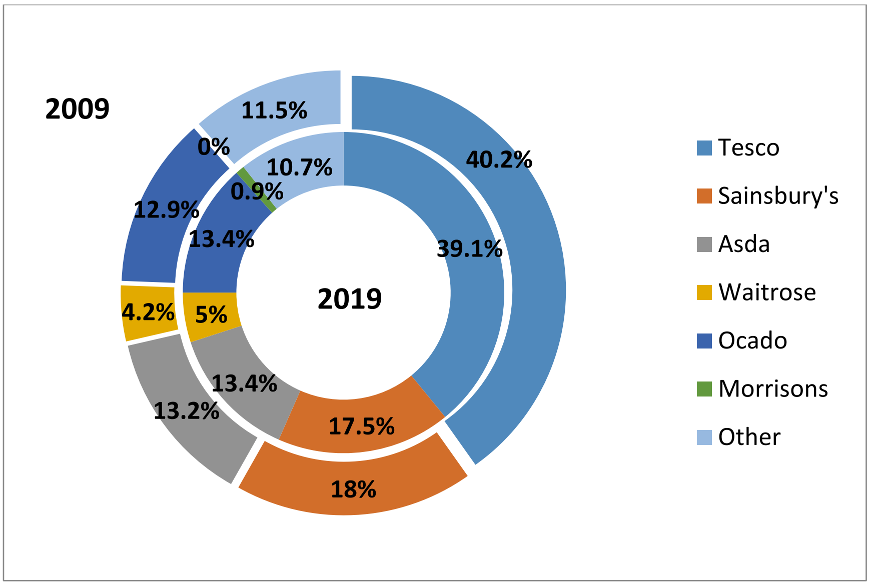


Figure 2.15: Online market shares 2009-2019. Source: Mintel 2015

According to Mintel researchers, online market shares will not change drastically by 2019 with Tesco staying as the market leader followed by Sainsburys', Asda and pure player Ocado (Mintel, 2015). However the pressure is on Tesco. For example, online only grocery retailer Ocado, despite its current low market share of 2.5% in total grocery market, now has 13% of the on-line market and has been voted as the best online supermarket every year in the "Which?" magazine opinion poll since 2010 and has enjoyed a massive 135% increase in its market share since 2006 (Digital Foodies, 2016; Ocado, 2016; Business Insider, 2016). The Ocado success is due to its innovative approach of centralised automated pick centres and customer focused retail environment. In addition, the fast developing online channel at Waitrose and the new market entrance of Morrisons are changing the distribution of online market shares since 2009, both predicted to seize more market share from Tesco by 2019. The pattern will change further with Amazon entering the market with its Amazon Fresh service in June 2016 (The Guardian, 2016).

According to data provided by IGD, approximately 20% of UK households buy their groceries online at least once a month, and 11% of UK consumers use online channels as their main mode for grocery shopping (IGD.com, 2016). Moreover, 20% of UK online users visit leading grocery retailers' websites every month, resulting in 3.5% or 1.3 million unique online grocery buyers every month (Kantar Media, 2012). According to IGD (2016), over a quarter of UK consumers buy their groceries online monthly compared to 22% in 2010.

As noted in section 2.7 above, consumers state the flexibility to shop any time and the convenience of groceries being delivered to their door (especially heavy items) as the main reasons of their choice of online channels (Kantar Media, 2012). Moreover, online shopping is more likely to be planned than spontaneous (compared to the in store experience) and e-shopping tends to be of a larger value, with 12% of all grocery transactions of £60 or more made online during one visit compared to 99% of all transactions worth less than £60 which are made in physical stores (Intelligent Positioning, 2013). According to IGD (2015) on average online shoppers spend £75 per visit, with over 50% of respondents claiming to spend between £51 and £100.

Figure 2.16 demonstrates the gradual increase in online expenditure on food and drink in the UK since 2007. According to this analysis, by market research company Mintel, the highest growth of online market share is within the period 2011 to 2016, with a 6-7% yearly increase,although growth is predicted to slow down after 2016 with annual increase down to 5%.

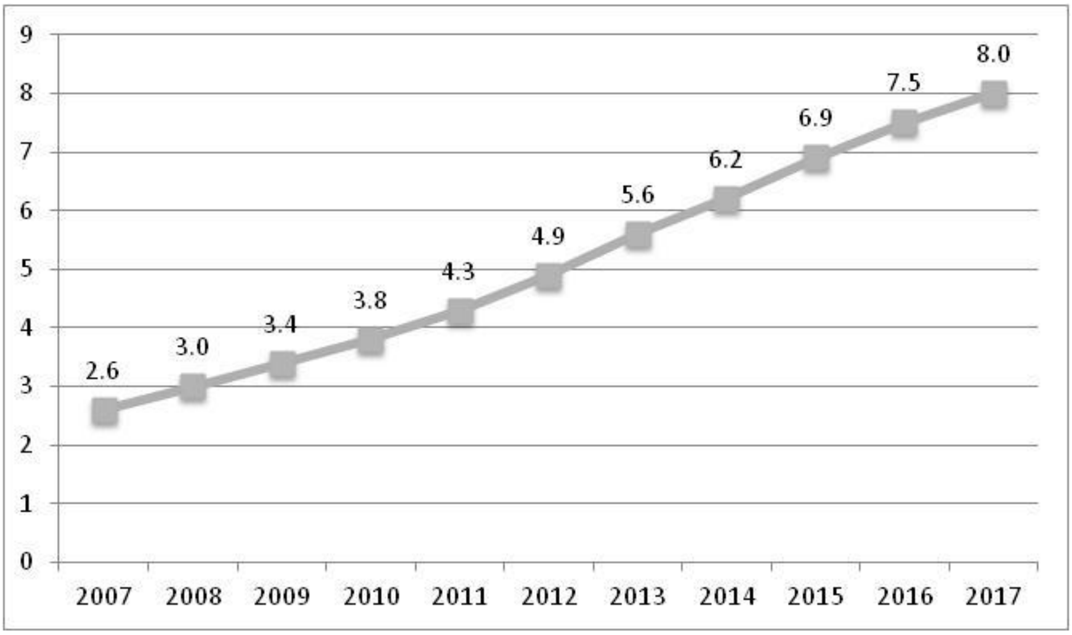


Figure 2.16. Online food and drink sales as % all spending on food and drink, 2007-17. Source: Mintel, 2013

These 2013 prediction rates were a little over optimistic, with online channel currently accounting for 6% in the total grocery sales which indicates a lower rate of new grocery channel adoption than predicted. The future growth of online sales predicted by Mintel and other research organisations is likely to be driven by the diffusion of new technologies and occurrence of new trends – m-commerce and t-commerce (shopping via mobile or tablet device discussed earlier in the chapter), which allow greater convenience with the ability to shop on the move (Mintel, 2013). During 2012 approximately 13% of the total online grocery shopping was completed using smart phones with total sales of £800 million (Mintel, 2013). According to Google, searches from mobile phone devices for on-line grocery doubled in 2014 and the Internet only based supermarket Ocado reported that almost half of all transactions were completed via a mobile phone device (Essential retail, 2014). However, technology is not the only factor which drives online sales.

So far, however, on-line sales have been limited demographically. For example, four in ten British consumers have never shopped on-line for their groceries and the number is highest among household of over 65s, those living alone and from the lower income (DE) socio demographic groups. The demographic distribution of online customers will be explored in the next Chapter.

To be successful in online grocery sector retailers need to provide cost-effective order fulfilment and overcome logistical issues of providing an efficient home delivery service (Fernie et al 2010). The challenges consist of picking the order (an average of 60 to 80 items)from three different temperature regimes from the wide spectre of products (between ten and twenty five thousand items) and delivering within 12 to 24 hours to customers home addresses at the particular time slot requested. For example, Tesco is delivering 250,000 online orders every week. The new innovative approach is required to service such a complicated and immense operation and existing models of delivering non-food products are not suitable in the long run despite the long history of mail catalogue shopping and the developed online channels of the major high street retailers. First, e-grocers require large number of vehicles dedicated to home delivery service. Secondly, the socio demographic profiles of online grocery customers differs from mail catalogue shoppers which creates new geographical patterns of home delivery with the widespread locations from rural areas with restricted accessibility to city centre locations with constant traffic congestions. Finally, customers have an immediate need for groceries and expect speedy delivery (Xing and Grant, 2006). Delivery is a very complex task for the retailers as it has to be secure, cost effective and satisfy customers' needs. Customers require convenience, reliability and punctual delivery. Problems can therefore arise easily. Unsecured delivery or 'door stepping' is when the driver simply leaves the grocery at the front door which may be more convenient for the customer but can have security issues from possible theft or damage (Fernie, 2010). Moreover, the fulfilment of online orders is far too expensive with supermarkets loosing approximately £300 million a year from online channel which is £3 to £5 loss on every order (The Financial Times, 2016). To minimise the shortfall grocery retailers are increasing the minimum online orders and introducing surcharges on top of delivery charges. For example, Tesco raised the minimum online order from £25 to £40 in 2015 following Asda's example earlier in the year. This move naturally created additional challenges in terms of encouraging existing customers to use online channels regularly and attract new online customers.

A possible way forward will be to strategically place grocery collection points around stores, petrol station and transport terminals which would tick all three requirements for effective delivery service – customer satisfaction, security and commercial viability. The latest strategy of supermarkets to encourage 'click and collect' service is proving to be popular with one in four of online shoppers using this channel in 2014 to buy their groceries, accounting for 5.2% of all online grocery orders (IGD, 2016b; Mintel, 2015). Currently most major supermarkets offer this format with Asda a leader with 400 stores across the UK offering 'click and collect' and 250 stores offering the same day collection service. Some of their stores have a 'drive thru' option also (Mrs Bargain Hunter, 2014). Specialists at Alvarez and Marcel (2014) are very sceptical about the profitability of online channels with the prediction that by 2018 only three out of six major supermarkets (Tesco, Asda, Morrisons, Sainsbury's, Waitrose and Ocado) will benefit from it unless they modify their methods or increase prices (despite the possible expansion of 'click and collect' services).

To sum up, the new online fulfilment model for on-line retailing should include the following elements (Alvarez and Marcel, 2014);

- Satisfaction of customer's needs of flexible deliveries during the days or various time slots;
- Widespread regional coverage to reach the target audience and increase market share;
- Optimisation of the online fulfilment costs and allowances for demand fluctuations (seasonality);
- Effective use of technologies to reduces expense and offer an excellent service. Currently, 55% of the total costs are incurred during the order picking stage;
- Application of different models depending on the location – local, national and regional. The model which works in very urbanised areas will not be suitable in rural localities;
- Profound knowledge of customer profiles regionally and nationally.

These fundamentals can be accomplished by applying the following techniques:

- Transport the online products to the region using existing vehicles;

- Utilise local facilities to run the vehicles delivering multiple orders throughout the day depending on the demand;
- Create picking operations in close proximity of existing stores and merge with ‘click and collect’ service;
- Apply innovations and technologies to improve picking rates;
- Offer choices to customers either to collect their order at the store or delivery to the home address using the small vehicles at the stores. To maximise the margins customers can order extra products at the facility using automated terminal or mobile device (tablet).

By applying this checklist, the cost per order can be reduced by almost 50% by 2018. On the contrary, the cost per order will increase to up to £25 per order if supermarkets continue to use existing methods either store based or warehouse based (Alvarez and Marcel, 2014).

2.9.3. Conclusions: The consumer needs and omni-channel grocery retailing

In the previous sections the different channels in grocery retailing were identified; supermarkets including hypermarkets, convenience stores, discounters, and online which includes click and collect service and home delivery options. Multi-channel retailing within the grocery sector follows the same stages of the shopping process discussed in earlier sections but with some distinctions (Lieke van Delft, 2013)

1. Stimulation. The choice of the channel will depend on the desired product. For example, if consumers require essential groceries, e.g. bread and milk, it is likely they will choose the c-store. Although, different channels are possible for the same products. For example, top-up or convenience shopping accounts for £26bn in supermarkets compared to £21.5bn at c-stores (Seely, 2012).
2. Search. Consumers use internal sources of information (previous experiences) or external – friends, media, Internet. The time taken depends on the type and value of the product. The higher the value the longer this stage will take. Grocery products do not require extensive search. This stage may involve an evaluation of other stages, e.g. delivery times and after sales service.

3. Purchase or action stage. The provider makes their choice and the product is purchased.
4. Delivery. The consumers face various options – take products from shop, home delivery, collect from pick up point or pick up later in the shop
5. After sales service. Returns, complaints. Consumers have offline, online choices to contact the retailers.

The main factors in the decision making process for choosing the channel are evaluation of risk, price, search, effort evaluation and delivery time. (Gong and Maddox, 2011). The choice of the channel will depend on product type and price, which remains the most important factor when choosing the store with 41% of UK consumers stating it as deciding reason (IGD, 2014).

In 2015 it was estimated that 58% of UK consumers use at least four different channels over a month to purchase their groceries (IGD, 2015). The choice of channel also depends on the emotional aspect of shopping. According to IGD (2015), half of respondents stated that they enjoy the shopping experience online, although three in ten become easily tired and fed up. It takes over 20min to complete the online grocery shopping experience for 62% of online customers. Retailers recognised the importance of speed and efficiency for online customers and introducing, for example, specialised express shopping lists. To maximise their profits, retailers need to make online shopping experience more enjoyable and encourage to spend more time online shopping and browsing which will urge impulse buying.

So far there was little said about variations in the geography of e-commerce sales. These variations will arise for a number of reasons, the most important of which are the variations in demand – not all consumers are equally likely to be e-commerce shoppers. The geodemographics of demand for e-commerce will be explored in the next chapter.

Chapter 3: E-retailing: customer demographics and spatial distribution of e-commerce shopping

3.1. Introduction

This chapter will examine the spatial distribution of Internet users and online buyers and the impact of spatial variables (such as location and shop accessibility) on e-commerce activities. This chapter will test two hypothesises: the innovation-diffusion theory versus the efficiency theory (see below). Furthermore, this chapter will empirically evaluate these theories in the current economic climate and make suggestions for possible other contributory factors to explain the geography of online sales. First, section 3.2 introduces the idea of customer segmentation and provides a description of demographic segmentation techniques applied in this research: section 3.2.1 introduces a single demographic variable – social class; sections 3.2.2 and 3.2.3 provide an analysis of different multivariable demographic classifications which are applied in this research, ACORN and OAC, and provide a comparative analysis; section 3.3 looks at the nature of Internet usage in the UK. Furthermore, the demographics of e-commerce will be discussed in section 3.4. Finally, section 3.5 examines the demographic profile of e-grocery shoppers and the spatial distribution of online grocery customers in the study area by various demographic characteristics.

3.2. Customer segmentation and demographic classifications

There is a long history of retail market segmentation research which is based on the assumption that customers have common features, interests, lifestyles etc. Market segmentation allows businesses to identify and target their potential audience to offer services and goods. There are various types of market segmentation based on customers' locations, personal traits and attributes. Segmentation by geographic location (regions, neighbourhoods, cities or postal sectors) is the most common and simplistic (Goldstein, 2007). Demographic segmentation is based on customers' characteristics, e.g. age, gender, level of education. Other segmentations (behavioural, psychographic and cultural) are concerned with consumer's lifestyles, activities, interests and cultural origin (ibid). The multivariable segmentation products combine these separate key characteristics. In this research the OAC and ACORN multivariable classifications were applied. The demographic patterns are complex and geodemographics analysis is one technique which enables representation of this complexity. Geodemographics is an established methodology which allows to capture multiple aspects of consumer behaviour within various areas. There are advantages of disadvantages using commercial demographic tools (ACORN) and openly available techniques (OAC). Commercial organisations such as CACI using the most advanced digital techniques and employ data not available for general public (CACI, 2016). Although, they widely apply national statistics data in the development of the their commercial products including ACORN and FRESCO. The commercial applications are very useful for organisations and projects which require tailor made products to find a solution to the particular problem. The publicly available demographic techniques are based on extensive survey and census data (ONS, 2016). Office for National Statistics uses the long established and constantly revised methodology with application of data linkage, harmonisations and the management and measurement of quality. The UK has a long history of producing an extensive free demographic classifications in comparison to the US where an openly available free classifications were not available up to almost twenty years (Singleton, 2014). Despite the existence of established demographic classifications there is a demand for a bespoke classifications as consumer behavior is becoming more complex with emergence of new social trends (e-shopping). In their study of e-society, Longley and Singleton (2009) developed application specific geodemographic classification to reflect the close connection of British society with emerging information and communication technologies (ICT). In this research both commercial and free demographic classifications are applied which are well established and for the purpose of this research it does not matter a huge amount a choice of typology. The detailed analysis of each market segmentation technique applied in this research is described in the next section.

3.2.1. Social class and other single demographic variables

The National Readership Survey (NRS) breaks down the population by social class and occupation. It was developed over fifty years ago and commonly used by researchers and marketers in various fields (NRS, 2016). There are six social grades from A to E starting with the households belonging to the upper middle class(with higher managerial occupations)and concluding with populations within the lowest social status, e.g. the lowest grade workers and pensioners. Table 3.1 provides a summary description of this classification.

Table 3.1. National Readership Survey (NRS) demographic categories

Social Grade	Social Status	Occupation
A	upper middle class	higher managerial, administrative or professional
B	middle class	intermediate managerial, administrative or professional
C1	lower middle class	supervisory or clerical, junior managerial, administrative or professional
C2	skilled working class	skilled manual workers
D	working class	semi and unskilled manual workers
E	those at lowest level of subsistence	state pensioners or widows (no other earner), casual or lowest grade workers

Source: NSR (2016)

Currently, the majority of the UK population is within grades B, C1 and C2 with 23%, 27% and 21% belonging to these groups respectively. The households within grades A and E are the least widespread with 4% and 9% correspondingly. Classification by social class is a very simplistic approach and doesn't consider behavioural and other demographic characteristics.

There are many other different variables which will affect human consumer behaviour. The most common and significant are age, gender, education, income, ethnicity, economic activity and family structure. The following 47 demographic variables have been analysed in this research in terms of estimating the demand for online expenditure in the study area (see Chapter 6).

Table 3.2. Demographic characteristics

Age Bands	Ethnicity	Economic Activity	Social Grade	Family Structure
Age 0-15	Total	Persons 16-74	Persons 16-64	Families
Age 16-19	White	Econ active	AB	Couple family
Age 20-24	Mixed	Employee	C1	Lone parent family
Age 25-44	Asian	Self employed	C2	Male lone parent family
Age 45-64	Black	Unemployed	D	Female lone parent family
Age 65+	Other ethnicity	Fulltime student econ active	E	Family 0 dependent kid
Male age 0-15		Econ inactive		Family 1 dependent kid 0-4
Male age 16-19		Retired		Family 1 dependent kid 5-18
Male age 20-24		Other econ inactive		Family 2+ dependent kids youngest 0-4
Male age 25-44				Family 2+ dependent kids youngest 5-18
Male age 45-64				Lone parent 1 dependent kid 0-4
Male age 65+				Lone parent 1 dependent kid 5-18
				Lone parent 2+ dependent kids youngest 0-4
				Lone parent 2+ dependent kids youngest 5-18

3.2.2. ACORN classification

In this study 'A Classification of Residential Neighbourhoods' (ACORN) demographic classification was applied to identify the profile of online consumers. This is the longest established commercial geodemographic classification in the UK (CACI, 2013). This customer segmentation technique categorises people into demographic types by postcode geography. ACORN also provides detailed classifications based on consumer's attitudes, multiple characteristics and lifestyles. This geodemographic segmentation was designed by CACI, a partner organisation for this project. This detailed classification includes 62 types, which are aggregated into 16 groups from A to P (Appendix A). The brief description of the six categories which are used in this research is provided below (CACI, 2013).

Affluent Achievers

The first category consists of the wealthiest households in the UK. They live in the prestigious rural, semi-rural and suburban areas of the country. They are usually aged mid-forties and older (retired wealthy pensioners), whose children have already left the household. Some neighbourhoods also contain large numbers of wealthy families with school age children, particularly the more suburban locations. These people live in large expensive houses, which are usually detached with four or more bedrooms. Moreover, these people are more likely to own a second property. The majority of Affluent Achievers are well educated and employed in managerial and professional occupations or have their own businesses. They have substantial income of over £60,000 a year with an index of 343 for earners of over £100,000 in comparison to national baseline figures (100). They are very technologically savvy and tend to use the Internet daily for managing finance, utilities and investments, searching for local and travel information, visiting lifestyle and weather websites. Moreover, they like to purchase online goods (beer and wine are the highest expenditures in this category with an index of 150 compared to the average national figures of 100).These people are healthy, wealthy and confident consumers. Figure 3.1 provides a summary of their lifestyles and interests.

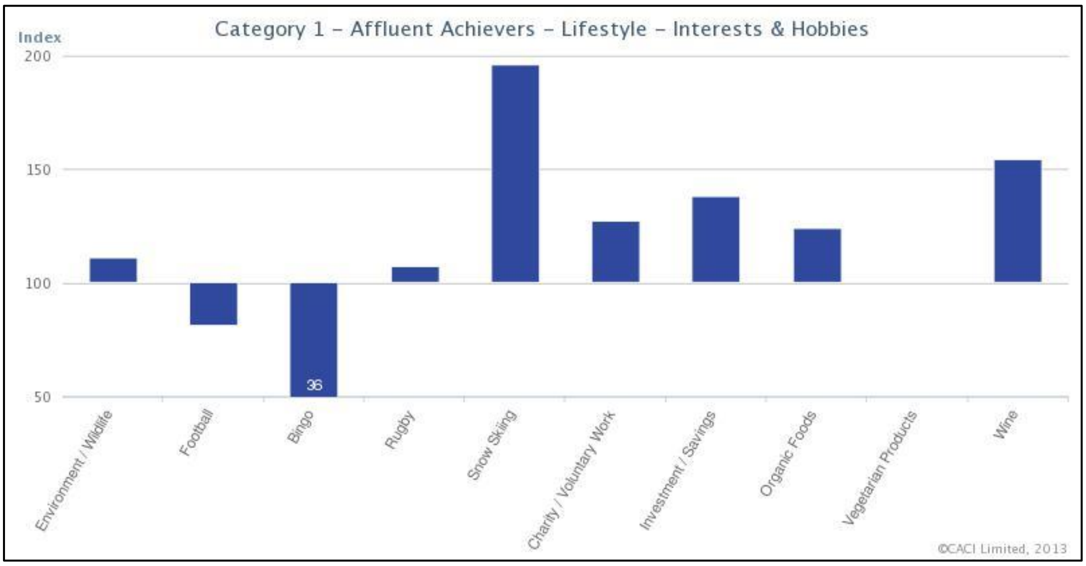


Figure 3.1. Affluent Achievers – Lifestyle. Source: CACI, 2013

These consumers are more interested in snow skiing, charity work and finances and least interested in bingo and football. They are the biggest spenders on wine and organic foods compared to their counterparts from the other five ACORN categories.

In terms of grocery shopping, Waitrose and Marks and Spencer are their favourite supermarkets with an index of 197 and 148 compared to the UK average of 100 (Figure 3.2). Asda and Morrisons are their least favourite grocery retailers. Finally, although enthusiastic Internet shoppers generally, Affluent Achievers are not the most enthusiastic online grocery customers compared to other ACORN categories with an index of 70 which is below national average (100). The profile of online customer will be discussed in more detail in the next sections.

Figure 3.2. Affluent Achievers – Preferences in terms of Grocery Shopping. Source: CACI, 2013

Rising Prosperity

These customers are well educated, cosmopolitans and mostly affluent young people living in urban areas. The majority of them are single or couples without children or with younger children. They are young professionals with successful careers. They live in modern executive apartments (although, some live in terraced townhouses). They are likely to rent their homes but a few have bought their homes. They are the most proficient technology users, an 'early adopters' generation and in possession of the new technological devices. They are outgoing, (often eating out) and enjoy entertainment (theatre, cinema and nightlife). Waitrose is their favourite supermarket and they are twice as likely to shop there as the average British consumer (Figure 3.3)

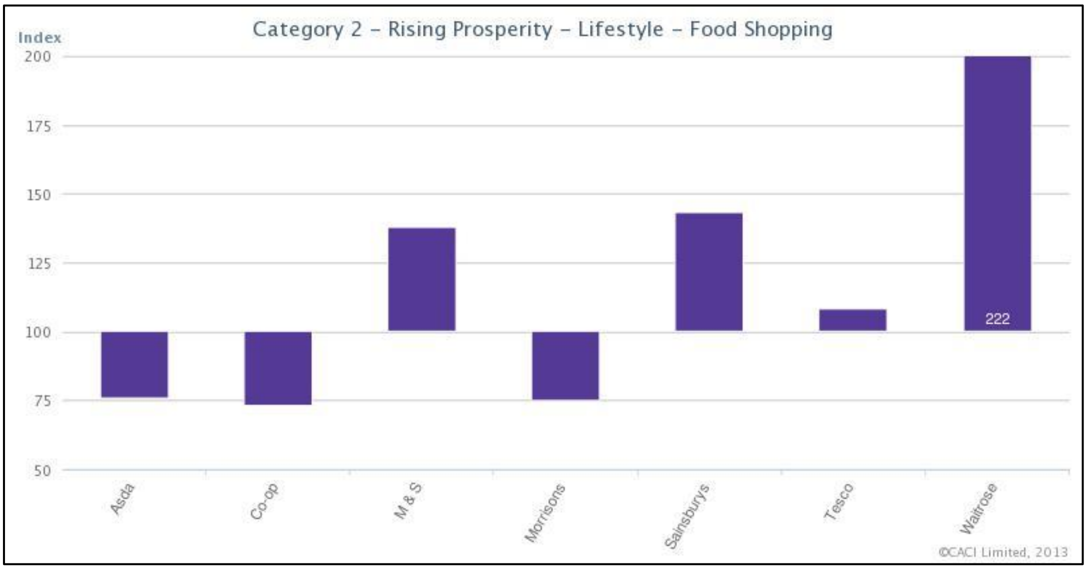


Figure 3.3. Rising Prosperity – Preferences in term of Grocery Shopping. Source: CACI, 2013

Their other favourite supermarkets are Sainsbury's and M&S with an index 143 and 138 compared to the national base (100).

Comfortable Communities

The households within this category can be described as 'average', whether young or older living in the suburbs, smaller towns or the countryside. They have families and live in suburban or semi-rural locations. They also include comfortably off pensioners, living in retirement areas around the coast or in the countryside and younger couples who have just started to live together. The majority of them own their home. Most houses are semi-detached or detached and overall, are of average value for the region. Incomes are also average with the younger people earning proportionally less. They might have some limited savings and investments. Their job occupations are within managerial, clerical and skilled fields. Educational qualifications are in line with the national average. They are not very wealthy but comfortable off. Figure 3.4 shows their grocery shopping preferences.

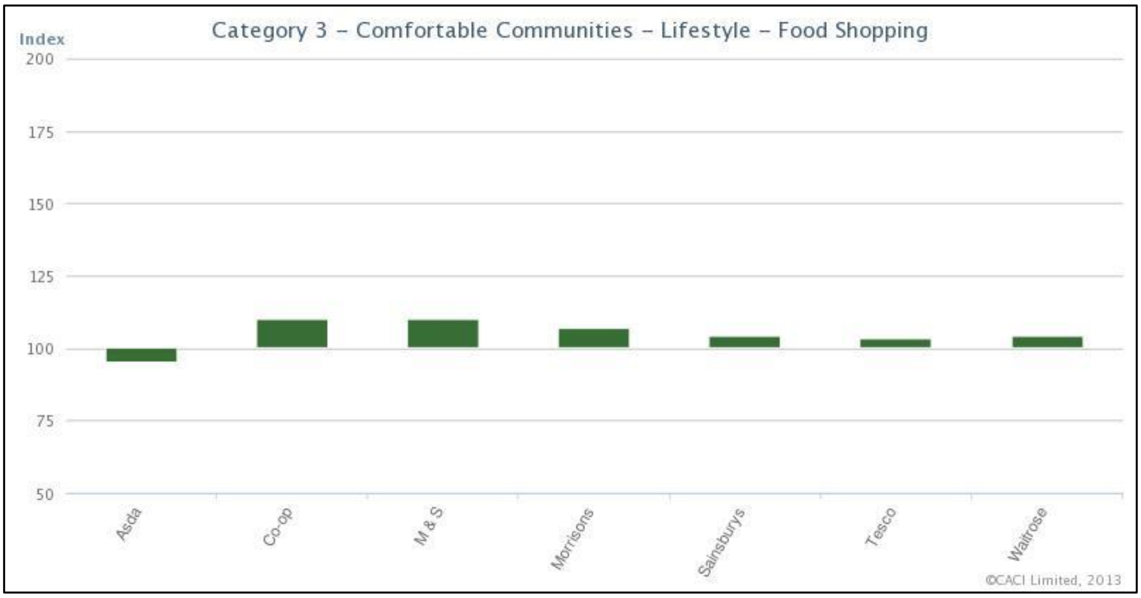


Figure 3.4. Comfortable Communities – Grocery Shopping. Source: CACI, 2013

Figure 3.4 shows that customers belonging to this category do not have substantial preferences for anyone particular supermarket, although, they are least likely to purchase their groceries at Asda with an index below national average. M&S have a slight advantage with an index of 110 compared to Tesco with an index of 103.

Financially stretched

Households within this category live in the traditional areas of Britain, usually in low value owner occupied terraced or semi-detached housing and council rented homes, including social housing developments specifically for the elderly. This category also includes student households. Incomes are inclined to be below average. They are engaged in lower paid administrative, clerical, semi-skilled and manual jobs. They tend to hold apprenticeships and O levels qualifications. The unemployment rate and benefit claimant rates are above average. These customers are less likely to possess a credit card, investments, saving accounts or participate in a pension scheme. They are more likely to have problems with debt and have been refused a credit card. These people are less likely than average to use new technology or to shop online or research using the Internet, although they will use the Internet for social networking.

Generally, the majority of households within this category have a modest lifestyle while some of them are experiencing financial difficulties. Figure 3.5 shows preferences in terms of grocery shopping.

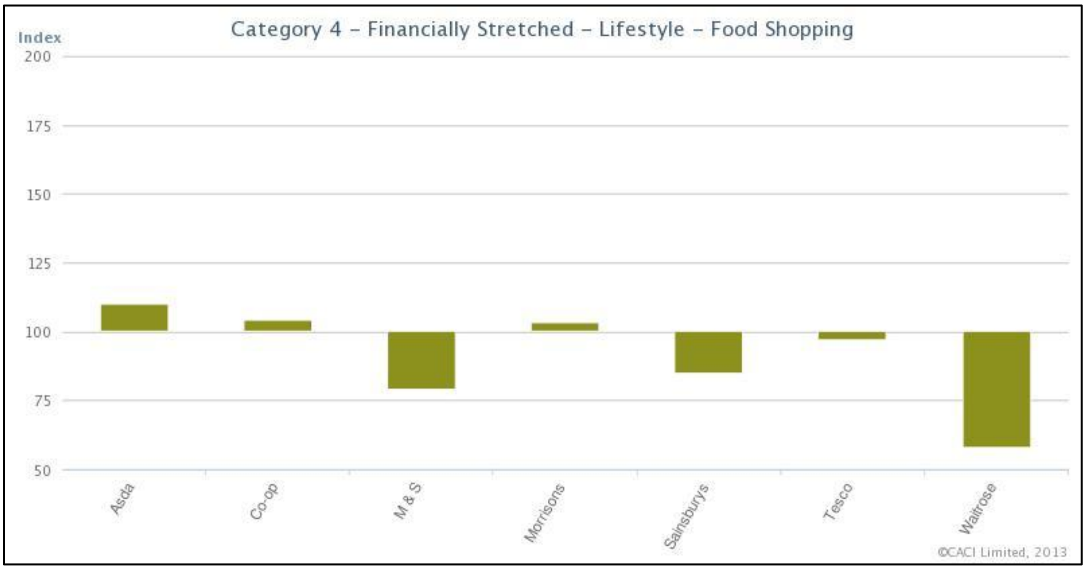


Figure 3.5. Financially Stretched – Preferences in terms of Grocery Shopping. Source: **CACI, 2013**

These customers naturally are least likely to buy their groceries at Waitrose, Sainsbury's and M&S (Figure3.5). Their favourite supermarket is Asda with slightly higher index of 110 compared to Co-op (104) and Morrisons (103).

Urban Adversity

This category contains the most deprived areas of towns and cities across the UK. Household incomes are low, nearly always below the national average. The number of people having difficulties with debt or having been refused credit is nearly double the national average. The numbers claiming Jobseeker's Allowance and other benefits is well above the national average. Levels of qualifications are low and those in work are likely to be employed in semi-skilled or unskilled occupations. They live in overcrowded terraced and semi-detached houses and purpose built flats, including high rise blocks which they are renting from the local council or a housing association. There is a small proportion of privately rented and owner occupier households. The households tend to be single adult, pensioners and lone parents households. As expected, customers in this category are not inclined to do their food shopping at the more affluent and expensive Waitrose with an index of only 49 compared to an average consumer. In fact, these consumers have preference for only one of the major supermarkets – Asda with an index of 115 (Figure 3.6). In addition, these customers are more likely to use discounters retailers (e.g. Aldi and Lidl), which are included on these graphs.

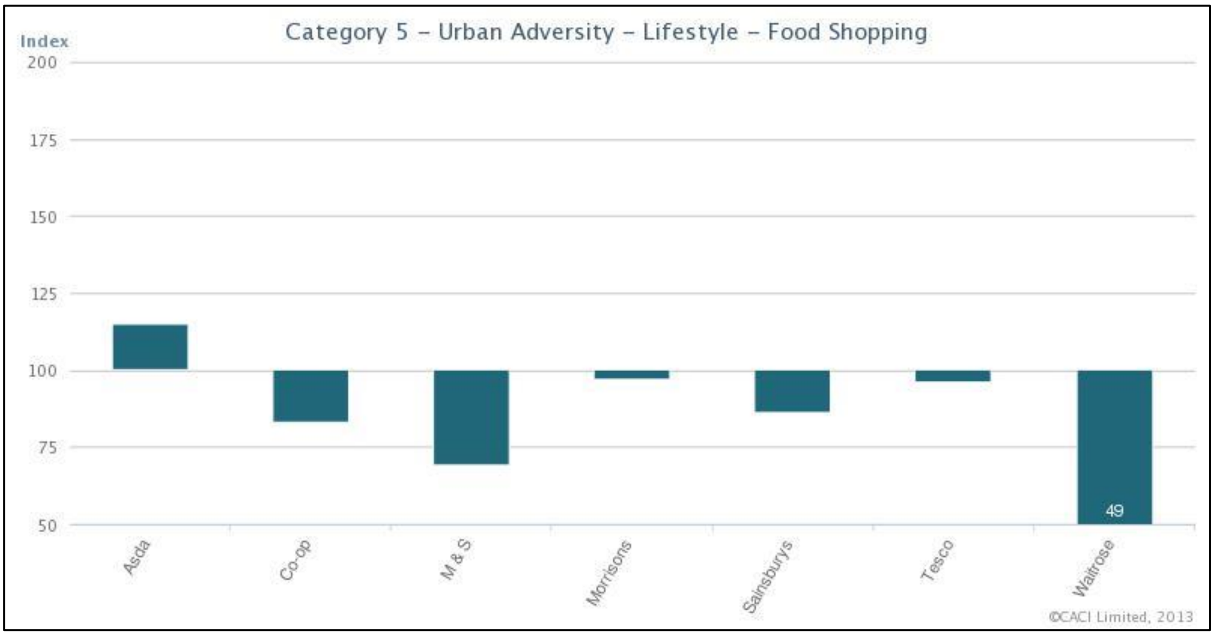


Figure 3.6. Urban Adversity – Preferences in terms of Grocery Shopping. Source: CACI, 2013

Not Private Households

These are the areas there the majority of residents are not living in private households and include military personnel bases, hotels, hostels, refuges, care homes and other communal council and medical accommodation. Although, some of these residents are potential grocery customers their contribution is minor and for that reason this category will not be included in this research. Moreover, some of these postcodes have already been incorporated into the previous five ACORN categories.

3.2.3. Output Area Classification

Another geodemographic segmentation technique based on multiple customer attributes is the Output Area Classification (OAC) which is used by ONS and was initially designed in 2001 by researchers at the University of Leeds and updated in 2011 by University College London (ONS, 2016b). This classification groups together geographic areas according to six key characteristics common to the population in these clusters – demographics, household composition, housing, socio-economic, employment and industry sector. The data is derived from the census data based on 41 variables with customers grouped into Supergroups, Groups and Subgroups (Vickers *et al*, 2005). Supergroups include seven categories – Blue Colour Communities, City Living, Countryside, Prospering Suburbs, Constrained by Circumstances, Typical Traits and Multicultural. The brief description of the seven Supergroups is provided below (Vickers et al, 2005).

1. Blue Colour Communities

Typically these are single, couple or lone parent households living in terraced housing working within construction, mining, manufacturing and retail industries. They belong to the white ethnic group and possess professional qualifications at the college level.

2. City Living

These customers are aged between 24 to 44, living in the highly urbanised areas in rented flats. They are highly educated and employed in managerial, financial and highly skilled jobs.

3. Countryside

These are older populations (45 to 64 years old and over) living in detached houses in rural areas. They have high car ownership levels with two or more cars per household. Many work within the agricultural sector and are likely to be self-employed and work from home. They also tend to do voluntary work.

4. Prospering Suburbs

Customers within this group are affluent and typically aged 45 to 64 living in detached housing with no dependent children, with two or more cars. They are of a white ethnicity, highly educated and are in managerial and professional occupations.

5. Constrained by Circumstances

These are the oldest population over 65 years old, pensioners' households or lone parent households living in social rented housing (flats), unemployed and/or with Limiting Long Term Illness (LLTI).

6. Typical Traits

As the name of this group suggests these are average UK households with no distinctive variables compared to the other Supergroups. They usually live in terraced housing, working part-time and belong to the middle age population group.

7. Multicultural

Customers in this group belong to various ethnic minorities mainly Indian, Pakistani or Bangladashi origin, who are born outside of the UK, living in crowded urban areas. They have high level of unemployment, use public transport and experience financial difficulties. A high proportion of students are within this category also.

There are five geographical levels at which OAC is produced. In this research the Output Area geography is applied. The advantages of OAC over other multivariable classifications include free access and wide practical applications, e.g. customer profiling and social marketing (ONS, 2016).

3.3. Overview of the development of Internet usage in the UK

Many scholars have indicated that there is a relationship between speed and quality of Internet connection and online purchase (Korgaonkar and Wolin 2002; Sexton 2002). This section examines UK Internet usage and the relationship with Internet access in particular.

UK Internet expansion became widespread due to the provision of unmetered Internet access after 2000, when users started to pay a monthly subscription fee instead of per kilobite (Mintel, 2013). The adoption of Internet facilities happened very quickly with the rate of Internet penetration increasing from 9% in 1998 to 86% in 2015 (Statista, 2016). The distribution of Internet access is represented in Figure 3.7

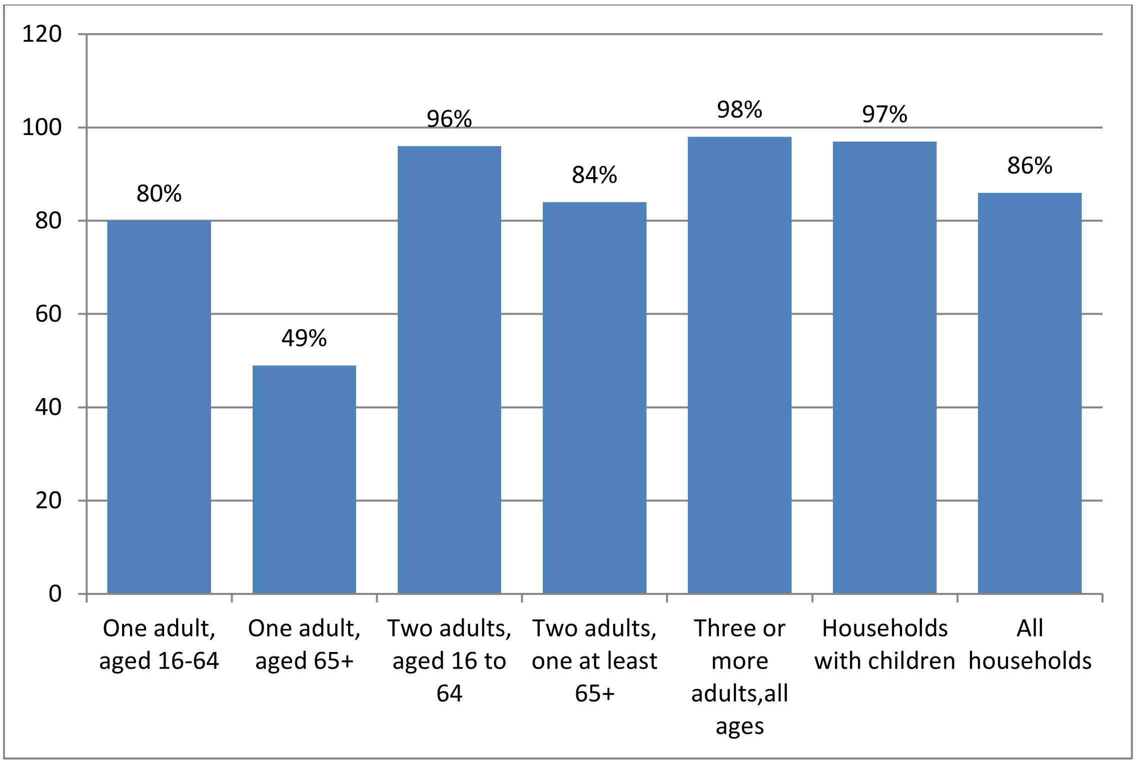


Figure 3.7. Share of the UK households with Internet access Source: Statista (2016)

Almost all households with children, with three or more adults of all ages and households with two adults aged 16-64 have Internet access compared to households with one adult aged over 65 where only half of households possess Internet access. In 2016, eight out of ten adults aged 16 and over(82%) used the Internet on a daily basis which is an increase of 57% in the last ten years (ONS, 2016). The individual's ability to use the Internet is linked to computer usage and access. In 2015 72% of the adult population in the UK used a computer regularly compared to 10% of adults who never used a computer (with only 1% of these adults aged between 16 to 24, ONS 2015). An analysis of the activities undertaken on the Internet by age group shows that young people (16 to 24) are more engaged in Internet recreational activities, e.g. social networking, watching TV and uploading content on the website (Figure 3.8)

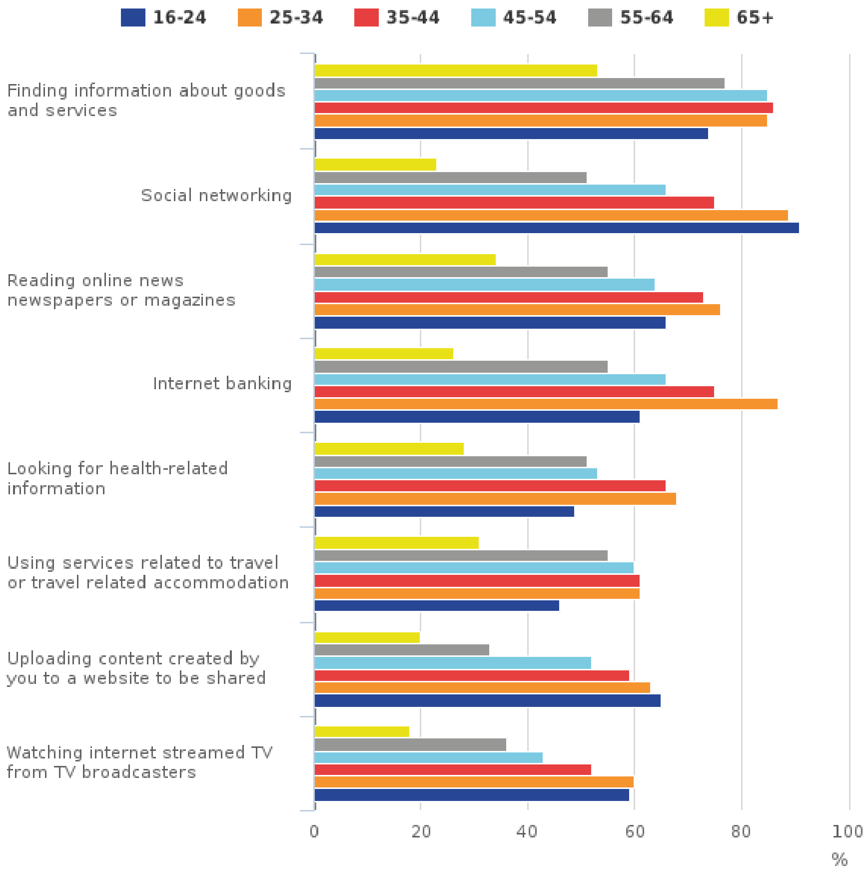


Figure 3.8. Internet activities by age. Source: ONS, 2016

Internet banking and looking for health related information are the most popular activities among adults aged 25 to 34. Finding information about goods and services are the most popular online activity among all age groups apart from young adults 16 to 24.

By 2016 almost all households with Internet access had a fixed broadband connection such as DSL, cable or optical fibre (ONS, 2016). One third of households now use mobile broadband via mobile phone networks as their Internet connection. Interestingly, in 2016, 59% of households without access to the Internet reported the reason as “they do not need it”. The other major reasons are the high costs associated with Internet usage and insufficient expertise to use the Internet (5% and 21% respectively). The other reasons for households not having access include privacy or security concerns and physical or sensory disabilities. In 2015, 27% of adults with disabilities had never used the Internet (ONS, 2015).According to European Commission researchers, in 2011 only 11% of the UK population had never used the Internet (European Travel Commission, 2013).

With regards to social class and income, almost all of the respondents (98%) with gross weekly earnings of £500 or more used the Internet, regularly compared with 7% of the respondents with weekly income of £200 or less, who have never used the Internet (ONS, 2012).

Analysing the above data, it can be seen that in the last 10 years the number of frequent Internet users has increased by over 50%, with the most active users being young people aged between 16-64, living in homes with 3 or more adults, or 2 adults aged 16 to 64 and families with children. Significantly, growth has also come from growing numbers of the older age group 65+, which indicates that the age gap between Internet users will narrow as present 50 and 60 year olds reach retirement age and will contribute to narrow with growing numbers in the last age category 65+ in the future.

Despite the significant growth of Internet usage, 11% all households still do not have access to the Internet due to a lack of desire in doing so, insufficient knowledge and skills or costs, and a lack of means to install it. The question arises as to whether this is due to a well informed choice or insufficient knowledge of the benefits and advantages (of what information technologies can deliver). The danger is that members of these households are detached from the benefits technologies can provide. Some commentators have consequently defined a “digital divide”. Longley, et al (2006) identified four major disadvantages and inequalities of limited access to the Internet – deficiency of human knowledge, difficulties in the labour market, missed opportunities as consumers and social exclusion (Longley et al, 2006). For example, people who are computer illiterate will be at a greater disadvantage when applying for a job (especially a skilled position) and they are likely to stay unemployed or obtain a lower paid position. As consumers, people with poor computer skills are at a greater disadvantage as they will not utilise the Internet to find the best deals, e.g. airline tickets, and as a result they will pay higher prices. Finally, individuals without Internet access will become marginalised from the communities as they do not participate in social network websites. Longley and et al (2006) identified 23 categories of online user within 8 groups based on the level of engagement with technologies among the UK population, ranging from complete non-adopters in Group A to e-professionals in Group H:

Group A : e-unengaged (elderly, technology as fantasy, mobile is the limit)

Group B : e-marginalised (mobile explorers, don't know what is Internet)

Group C: becoming engaged (e-bookers and communicators)

Group D: e for entertainment and shopping type (light on line shoppers)

Group E: e-independents (learners, light users, rational utilitarians)

Group F: instrumental e-users (computer magazine readers, online purchases, e-exploring for fun)

Group G: e-business users (electronic orderers)

Group H: e-experts (e-committed and professionals) Source: Longley, et al, 2006

Figure 3.9 demonstrates Longley's et al findings and their classification of the UK population within these eight e-clusters. Evidently, at the time of the research in 2005, almost 40% of the respondents were considered as non-adopters and only 3% identified themselves as e-experts. Interestingly, within the groups, the genders were almost equally divided, except group G with males almost three times more proficient users of new information and communication technologies (NICTs) for business purposes than females. Regarding the age distribution within the clusters, unsurprisingly, the older population 55-65 and 66+ are the main contributors to Group A (31% and 46% correspondingly). The younger age groups 18-25, 26-35, 36-45 were the main contributors and equally distributed in clusters B, C and D. The majority of e-professionals in group H were young people aged 18 to 35. Geographically, the concentration of proficient Internet users in categories H and G were in the south, especially around the London area and around other large cities. Longley's et al findings confirm Roger's Diffusion of Innovations theory, stating that technology adoption begins by younger persons and in the cities (Roger, 2003). Obviously, the modern e-society classification will be very different with a much lower percentage in category A. Nevertheless, this model is useful in the classification of Internet users and identifying reasons behind the "digital divide".

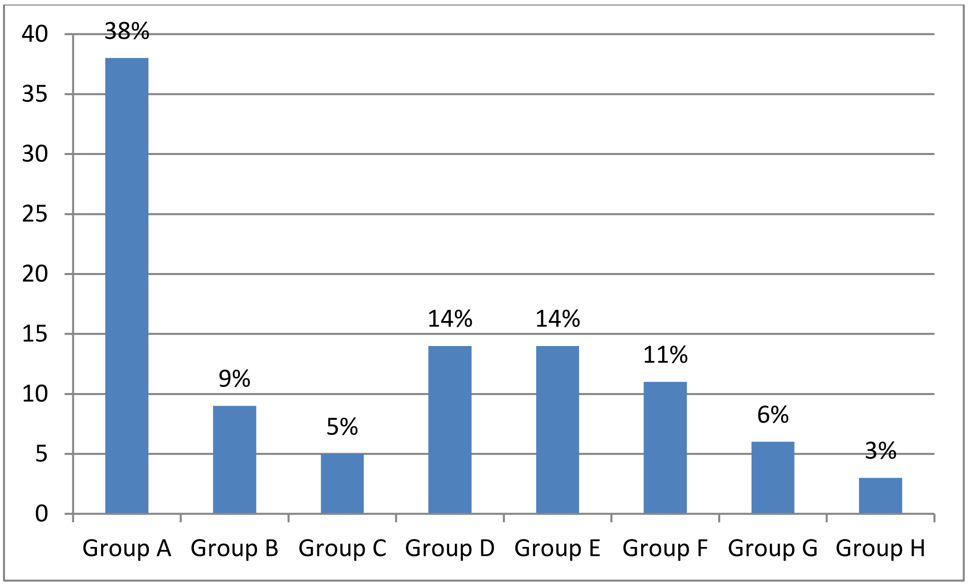


Figure 3.9. Classification of e-society. Source: Longley et al, 2006

Longley et al conclude that there is little empirical research about the spatial differentiation of "digital divides". The Axciom ROP data of broadband access distribution across the UK indicates that in 2010 the south and east of England had a greater advantage of broadband connection compared with rural areas, or Wales, the north of England and Scotland (Clarke et al, 2015). However, since then the UK Government has identified areas with slow broadband connections and assigned £363m investment into fast speed broadband of more than 24 megabits per second in all areas of "not-spots" and "not-a-lot spots", and narrowed the gap between rural and urban communities (Bradshaw, 2011). In 2016 there was a 10% difference between the highest Internet access rates in London and the South East with 94% of households with Internet access, and the lowest access in the West Midlands. 86% of households in the study area of Yorkshire and Humberside have access to the Internet which is below the national average figure of 89% (ONS, 2016). According to The Tinder Foundation (2015) households with no access to Internet are losing an estimated of £560 a year from not shopping and paying bills online.

3.4. Geodemographics of e-commerce usage

Many studies have indicated the variance in online buying among different demographic groups. In the US in the late 1990s the usage of online retailing was mainly by men aged between 25 to 45, college educated with high incomes (Mintel, 2013). UK Internet users followed suit and it is accepted that young males are early adopters of new technologies. Nowadays the demographics of Internet shoppers reflect more the general distribution of the population, i.e. women are becoming very enthusiastic online users.

In the last six years the growth of online shopping has been mostly driven by consumers aged 55 to 64 with a 33% increase in online purchasing within this group since 2008 (ONS, 2016). In addition, almost half of the older population aged 65 and over completed a purchase online in the last year (2015) which is an increase of 30% compared to 2008. That said, the most enthusiastic online shoppers still belong to the younger age groups, with 9 in 10 customers aged under 55 purchasing online in the last 12 months. Figure 3.10 shows the popularity of online products among various consumer age groups. Clothes and sports goods were the most popular category in 2016 with over half of adults purchasing them online, with younger customers aged 25 to 34 buying these products more frequently (73% compared to 24% of customers aged 65 and over who purchased these products online). Household goods are the next popular items with almost half of the adult population aged 16 and over purchasing them online in 2016. In 9 out of the 14 categories, customers within the 25 to 34 age group scored the highest in each category. Computer hardware was the least popular online category among all age groups, followed by video games, software and upgrades which are most popular among online shoppers aged 16 to 24 (who are the leaders in this category with 36% of them buying these products online in the last year).

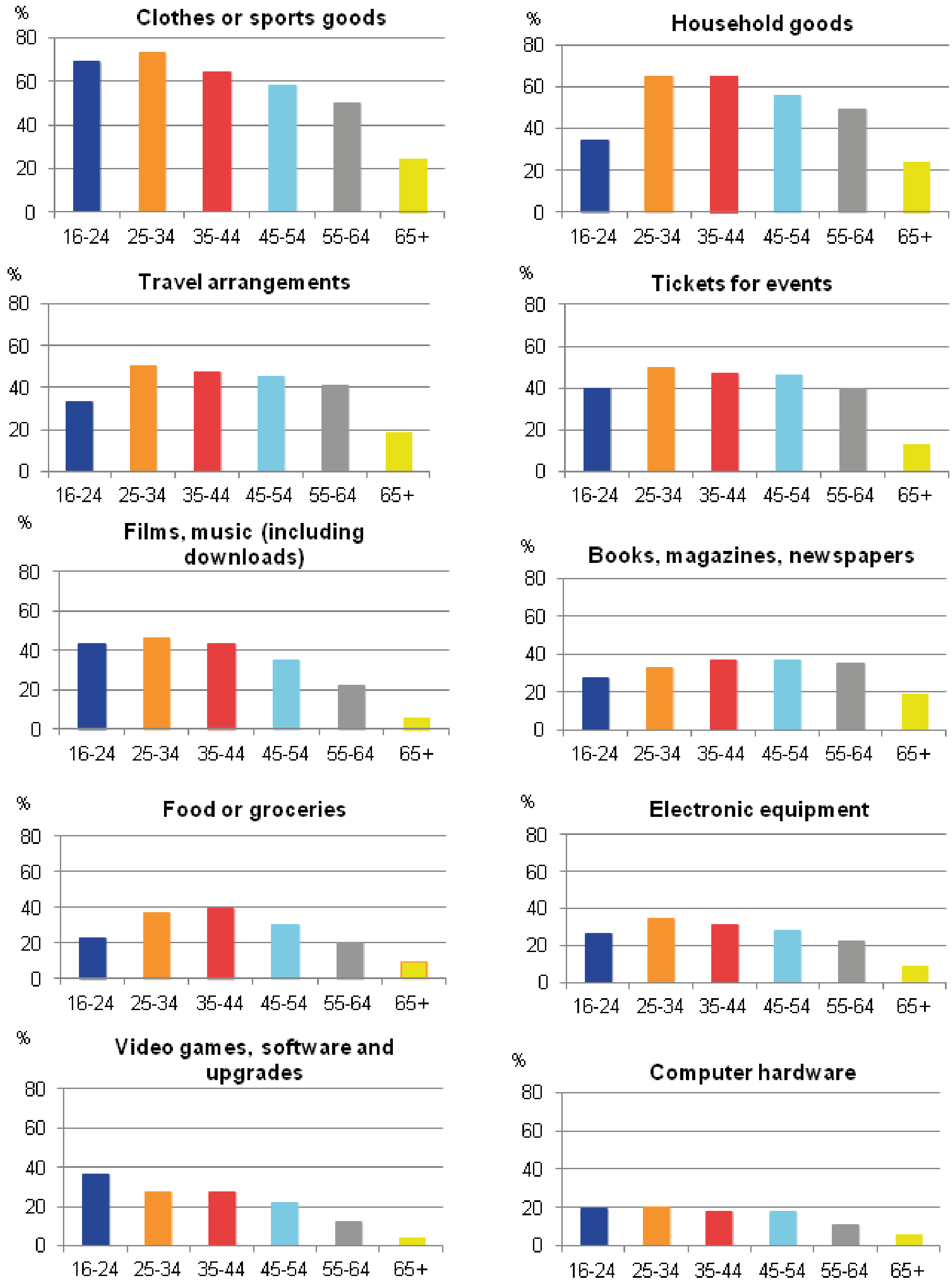


Figure 3.10. Online purchases by age groups and product type, 2016. Source: ONS, 2016

The most frequent online customers with 11 or more online purchases over three months are customers within 35 to 44 age group.

What influences online customers to buy products or services at particular websites? The customer can be driven by offers from the retailers, product reviews or price comparison results (see longer discussion in Chapter 2). In terms of age, there is not much variation between customers of all ages in making a decision of online purchase induced by retailers, producers or service provider websites, with approximately 50% of them almost always gaining knowledge from these sources of information (Figure 3.11). Customers between 25 to 44 years old are more likely to purchase online with 53% of them almost always buying online as a result of reviews posted by others. The price and product comparison websites have lesser impact on online customer decision making processes when purchasing online compared with reviews by other. For example, almost half of young people 16 to 24 rarely or never check price comparison websites to complete online purchases compared to 34% of people aged 25 to 34, which is the highest number of people of one particular age group which is influenced by price or product comparison website results.

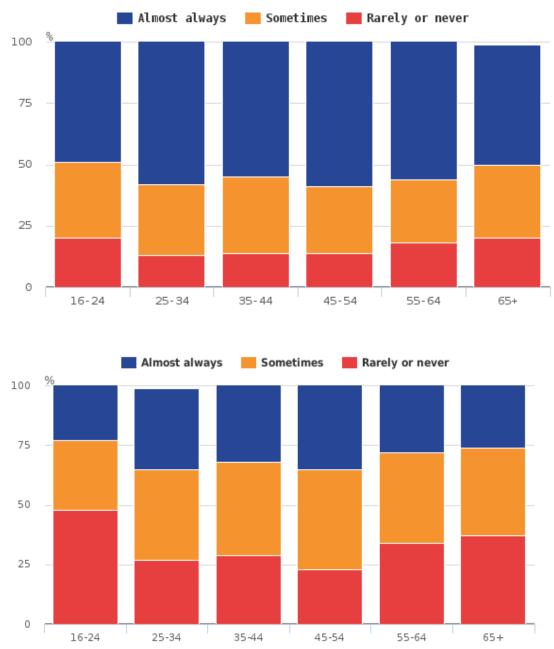


Figure 3.11.Use of information from retailers', producers' or service providers' websites and use price or product comparison websites or apps. Source: ONS, 2016

The variations associated with online shopping can be explained by two theories – efficiency theory and diffusion of innovation. The efficiency theory suggests that consumers living in rural locations with limited access to shops are more likely to shop online. The latter states that new technologies emerge in the cities and are initially adopted by young professional affluent males (Rogers, 2003). These ideas will be explored in more detail below.

Many scholars have examined the variation of e-commerce usage among different demographic groups (Weltereven 2007, Soopramanien and Robertson2007, Stroud 2009, ONS 2012, Clarke et al 2015). As noted in the introduction, most studies have been based on samples of consumers rather than actual consumer purchase data provided by retailers. This section summarises the key findings in relation to geodemographics in the literature which tend to support and confirm the findings discussed above in relation to the ONS (2016) report. The first variable which has emerged in the literature as important, and is clearly supported by ONS (2016) is the age of the consumer. Clarke et al (2015), exploring e-commerce consumers based on the extensive UK Acxiom Research Opinion Poll data, demonstrate that almost one third of all respondents aged between 25 and 44 are frequent e-shoppers, with the least online buyers in the age category 65+, with less than 10% of them frequently shopping on the Internet (Figure 3.12).

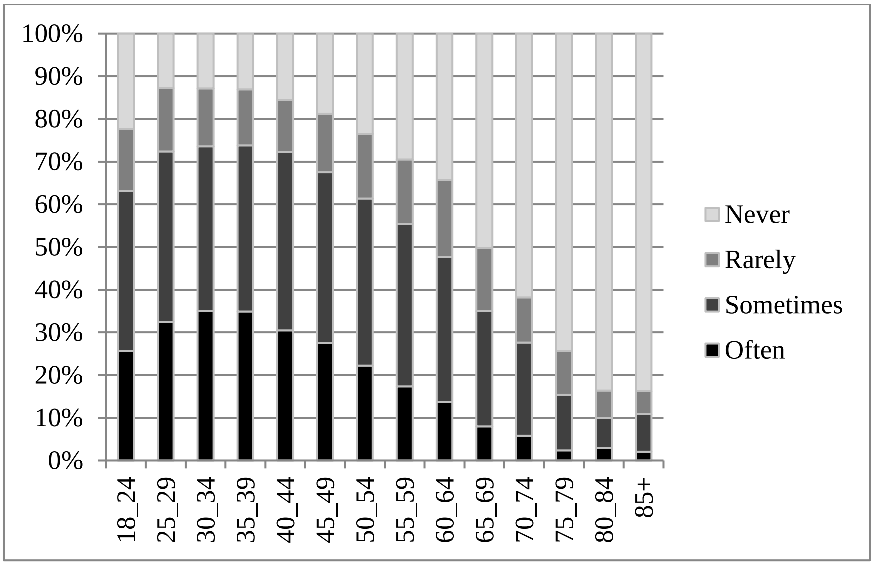


Figure 3.12. Percentage of households "How often do you use the Internet to buy goods and services" by age and gender. Source: Clarke, Thompson, Birkin, 2015

Interestingly, other surveys have found less connection between age and Internet usage. Stroud (2009) for example, argues that other factors are more important, with education argued to be the most important single factor (since, according to Stroud, 93% of the UK population under 70 with a university degree access Internet from home) followed by socio-economic factors AB and C1 classification groups (the most affluent professional social groups in the UK). However Stroud does admit that there is a sharp drop in Internet usage amongst the over 70 age group which relates to insufficient Internet experience prior to retirement. He argues that this gap between the over 70s and the younger age groups will disappear in the near future as the Internet experience will be more extensive for generations to follow.

It is useful to explore education or social class further. Other studies have shown the shoppers from the higher social class are more enthusiastic online customers with three quarters of the e-shoppers belonging to the AB social group (Evolution Insights, 2010). Researchers from Kantar Worldpanel established wealthier households are more frequent online spenders. Consumers with an annual income of £60k or more are spending 10% of their grocery budget online (Intelligent Positioning, 2013). Clarke et al (2015) also established that the wealthiest households are ten times more likely to buy online than the households on a lower income (Figure 3.13)

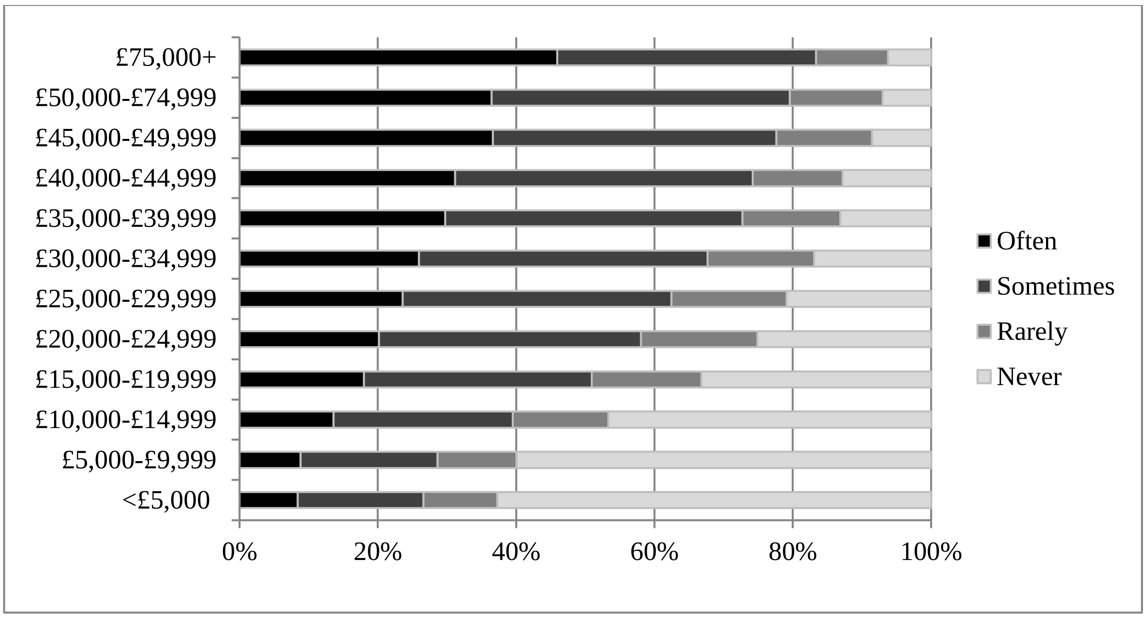


Figure 3.13. Variations in usage for "How often do you use the Internet to buy goods and services" by income: Source: Clarke, Thompson, Birkin, 2015

Given the fact that the focus in the thesis is on groceries it is useful to examine this in more detail in the next section

3.5. Profile of e-grocery customers

Table 3.3 summarises the demographic profile of the typical UK online grocery shopper and demonstrates the dynamics of online grocery purchases for the period 2011 to 2013. The data is provided by Eurostat, the statistical office of the European Union and based on the National Statistics Opinions Survey from 67 UK postal sectors with approximately 3000 participants each year.

Table 3.3. Distribution of online grocery purchases among various demographic UK population groups

Online purchases: Food/groceries	2011	2012	2013
individuals 16-24,	11	10	15
individuals 25-34	31	30	35
individuals 35-44	26	29	35
individuals 45-54	18	20	22
individuals 55-64	14	11	14
individuals 65-74	6	6	8
individuals 25-54 with medium formal education	22	22	26
individuals 25-54 with high formal education	33	35	39
individuals 25-64 with high formal education	30	32	35
individuals 25-64 who are employed	25	24	29
individuals 25-64 who are unemployed	20	n/a	16
males 16-74	16	16	20
females 16-74	21	22	25
males 25-54	20	22	26
females 25-54	29	30	34
individuals with no or low formal education	5	6	7
individuals with medium formal education	17	16	20
individuals with high formal education	28	29	32
individuals who are born in non-EU country	n/a	16	18
individuals who are foreign-born	n/a	15	20
individuals who are native born	n/a	20	23
Non-nationals	n/a	15	22
Nationals	n/a	19	23
Active labour force(employed and unemployed)	23	22	27
Individuals living in densely-populated area (at least 500 inhabitants/Km²)	18	18	n/a
Individuals living in intermediate urbanized area (between 100 and 499 inhabitants/Km²)	17	20	n/a
Individuals living in sparsely populated area (less than 100 inhabitants/Km²)	21	23	n/a

Individuals living in a household with broadband access	21	21	24
ICT professionals	n/a	40	n/a
Non ICT professionals	24	22	28
Non-manual including the armed forces	28	27	32
Manual	11	12	16
Retired and other inactive	10	13	13
Employees, self-employed, family workers	24	23	28
Students	10	11	14
Unemployed	17	n/a	16

Source: Eurostat, 2013

Again it can be seen that the younger populations are more enthusiastic online buyers with almost half of the UK population aged 16 to 34 buying their groceries online in the three months period. In addition, the greater increase of online purchases from 2011 to 2013 (20%) is within this age category, although there was a steady increase of buying grocery online among the other age groups. The other significant factor when establishing characteristics of online grocery shopper is again level of education. Almost 40% of individuals with high level of formal education within the age range 25 to 54 are more likely to be interested in e-grocery shopping and this trend is persistent across all age groups. In addition, over one of third of the UK population with higher education bought their food online during a three month period compared with only 7% of individuals with no or low formal education. This tendency is supported by the fact that 30% of individuals who are employed also buy their grocery online (as educated people are more likely to be employed). This survey demonstrates that females are more likely to shop online with the most enthusiastic shoppers within the age group 25 to 54, with a third of the participants undertaking their shopping online. Researchers from the Evolution Insights, Research Consultancy, arrived at the same conclusions regarding the age and gender of the typical online grocery shopper. They established that primary online shoppers are females aged between 25 and 54 (Evolution Insights, 2010). Figure 3.14 demonstrates the distribution of online shoppers with a quarter of online shoppers being aged 35 to 44 with 55% of them being females. Although, the distribution of online shoppers reflects the overall demographic profile of the UK adult population.

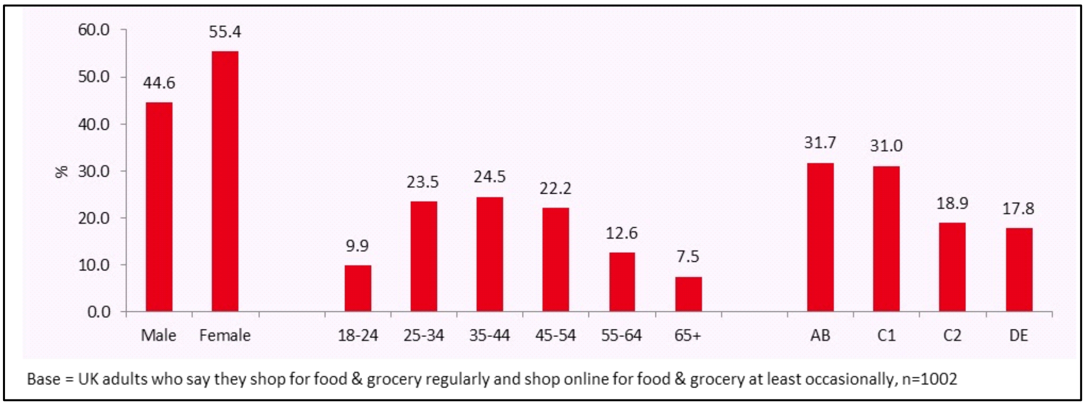


Figure 3.14. Demographic profile of online food and grocery shoppers. Source: Evolution Insights, 2010.

At the same time Figure 3.15 demonstrates considerate variation between market the penetration of online grocery shoppers by demographics. For example, females are much more likely to shop online with 70% compared to 58% of male shoppers and young people 18 to 24 are as likely to shop online as middle aged customers 45 to 54.

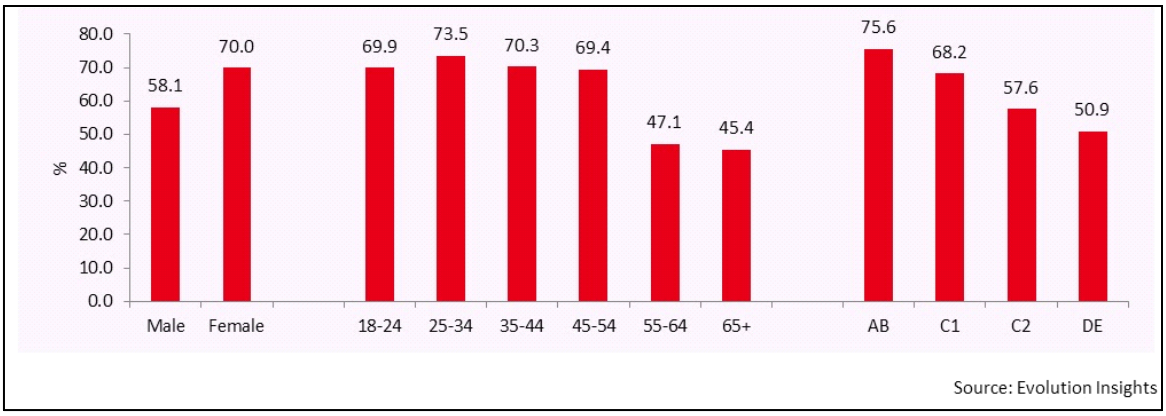
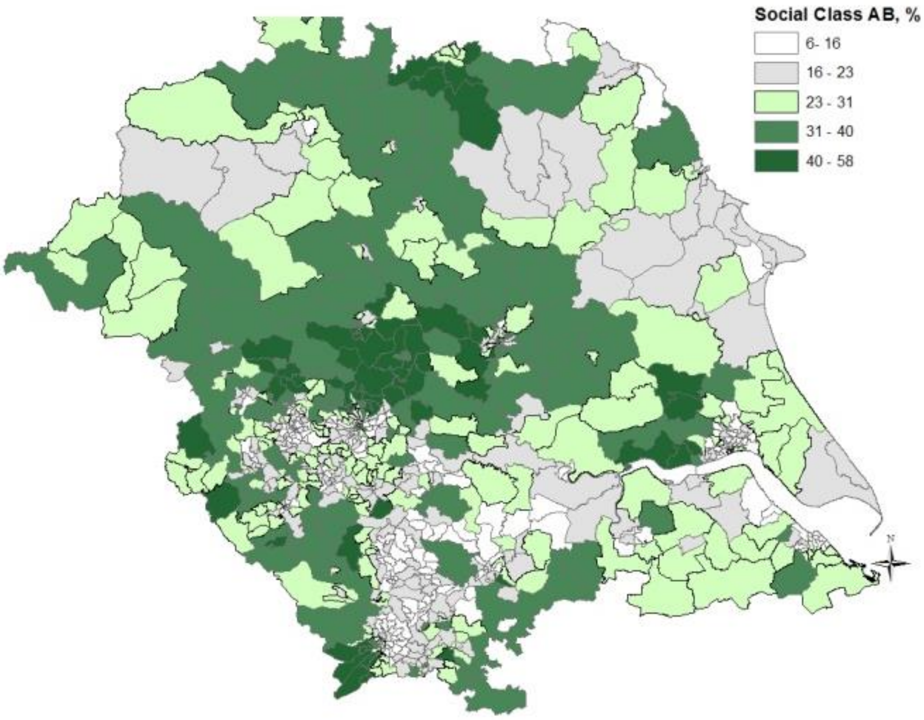


Figure 3.15. Penetration of online food and grocery shoppers. Source: Evolution Insights, 2010.

In addition, the shoppers from the higher social class are more enthusiastic online grocery customers with three quarters of the e-shoppers belong to AB social group. Researchers from Kantar Worldpanel also established that the wealthier households are more frequent online spenders. Consumers with an annual income of £60k or more are spending 10% of their grocery budget online (Intelligent Positioning, 2013), although researchers from the Institute of Grocery Distribution argue that consumers from contrasting socio-economic groups are equally interested in e-shopping with 20%

consumers from ABC1 and C2DE groups purchasing their groceries online in the last month (IGD, 2015). The data from Eurostat (Table 3.3) indicates that professionals and non-manual workers are twice as likely to shop online compared with manual workers. Moreover, ICT professionals are the primary demographic category for online shopping with 40% of them purchasing groceries online in the three months period in 2012 (Table 3.3).



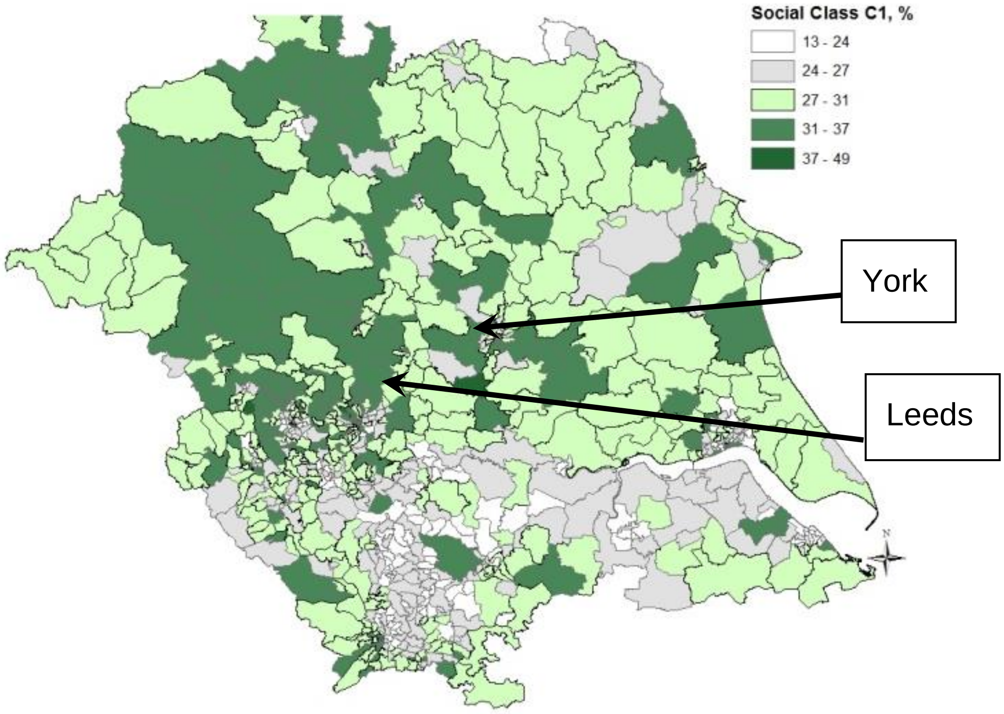


Figure 3.16. Social Class AB, C1 – Yorkshire and Humberside

Figure 3.16 shows the distribution of population in the study area by AB and C1 social class, which are more likely to be online grocery shoppers as has been demonstrated above. Based on these demographic characteristics the most likely locations of e-grocery customers are in the rural and suburban areas in the central part of Yorkshire and Humberside and around more affluent suburbs in the large cities, e.g. Leeds and York (this will be tested in Chapter 6 when the distribution of e-commerce users is formally estimated).

Thus, so far, the implication is that Internet usage will be higher in areas of high income and where there is a substantially younger population. Again Clarke et al (2015) highlight this through their map of the distribution of e-commerce users in Leeds, data again based on the Acxiom consumer survey (Figure 3.17). Area A on Figure 3.17 shows the higher usage in the northern suburbs of the City which are the most affluent. In contrast, area B on Figure 3.17 shows the less affluent south and eastern suburbs having little or no regular e-commerce activity in the grocery market. Area C on figure 3.17 shows the student area of Leeds – again high Internet usage as it can be expected (even if transaction value might be low). Area D begins to hint at other important issues. Although the area is relatively affluent, it is also increasingly rural. This starts to back up the ‘efficiency’ hypothesis of other studies - that consumers living in rural locations with limited access to shops are more likely to shop online (Farag, 2006).

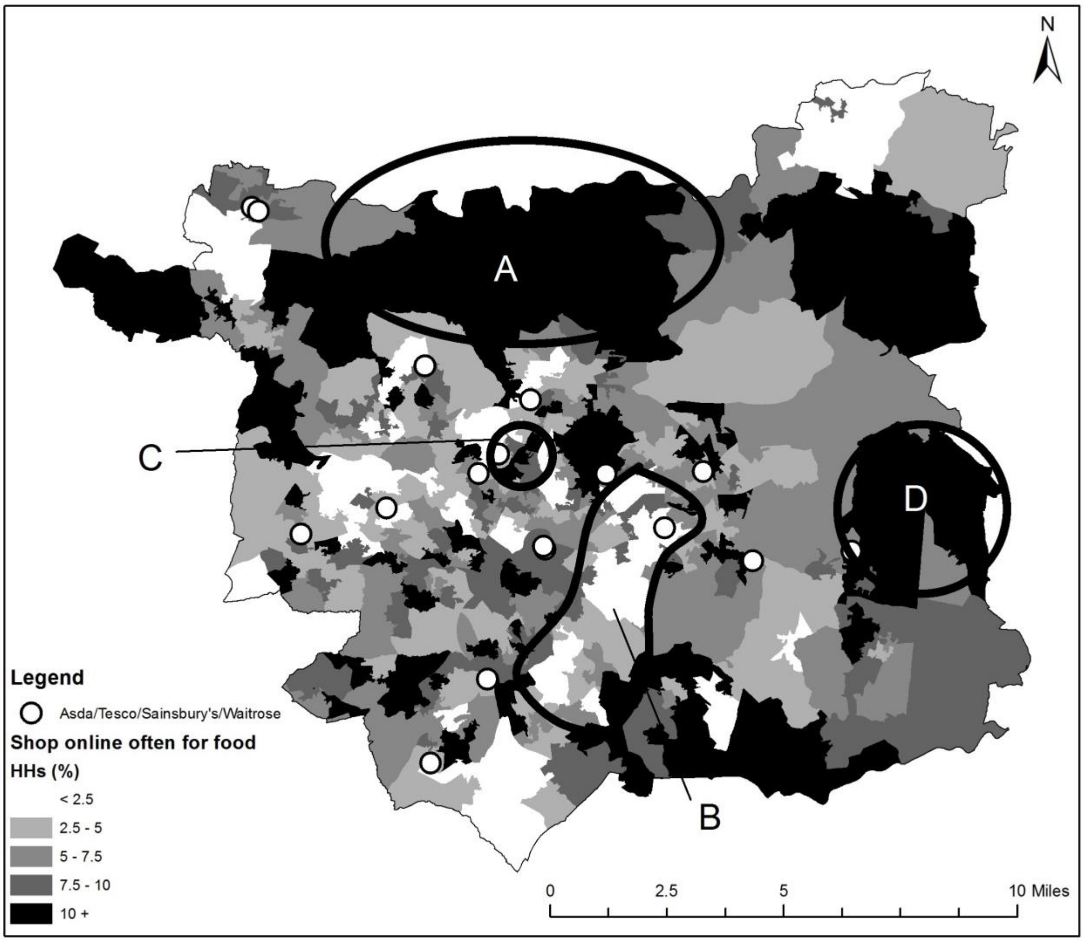


Figure 3.17. Online grocery penetration by lower Super Output Area in Leeds based on Acxiom ROP data. Source: Clarke et al 2015

If the characteristics of e-commerce shoppers are indeed multi-faceted, then it is useful to explore the potential of geodemographic systems to enhance the analysis presented thus far. As noted earlier, Longley et al (2006) identified 23 categories of online users within 8 principal groups (using cluster analysis) based on the level of engagement with technologies among the UK population, ranging from complete non-adopters of e-commerce in Group A to e-professionals in Group H. Similarly, a study by the Office of National Statistics in 2010 (the British Population Survey) explored the combination of various socio-economic characteristics in relation to online expenditure using the UK Output Area Classification (OAC) system. This is an open-access geodemographic system built by academics for the UK Office of National

Statistics. Figure 3.18 indicates the preferences towards e-grocery shopping among the different OAC supergroups compared to the national average figures derived from the British Population Survey based on the information collected from over 80000 individuals every year (ONS, 2010).

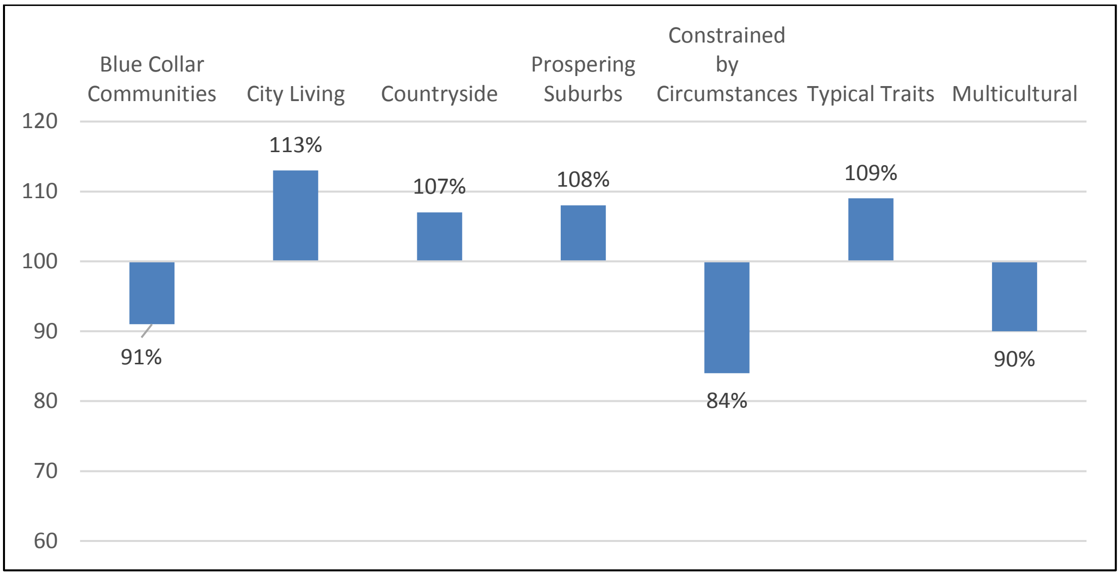


Figure 3.18. Online grocery preferences by OAC groups: Source: ONS, British Population Survey, 2010

According to the BPS data there is a clear indication that Blue Collar Communities, Constrained by Circumstances and Multicultural communities (the least affluent groups) are not enthusiastic online shoppers. However, within these groups there are distinctive variations with some Older Blue Collar and Afro-Caribbean communities more interested in e-commerce(with the scores above average). Despite the fact that all subgroups scored above average in the City Living category, their results were not particularly high compared to the Accessible Countryside subgroup, where 30% are more likely to buy groceries online compared to the average UK consumer. This fact again supports the efficiency theory stating that less accessible areas will be greater e-commerce users. The other enthusiastic online users are Prospering Suburbs and Typical Traits.

The multivariate analysis of demographic profile of online shoppers represented in the Table 3.4 supports this analysis with Prospering young families and Aspiring Households being the principal online shoppers among other Output Area Classification categories. Interestingly, Table 3.4 demonstrates that there is no correlation between frequency of Internet access and e-grocery buying. There is a quite opposite tendency in fact with Aspiring Households scoring below average in Internet access category and Transient Communities, the least frequent Internet users, being5% more likely to shop online than an average UK consumer. Moreover, agricultural communities are at the bottom of the table for online shopping, although they are the most frequent Internet users.

Table 3.4. Frequency of Internet access and online grocery shopping by OAC

Output Area Classification	Internet Access: once a day	Internet Access: once a week	Internet Access: 2/3 times a month	Grocery shopping in the last 3 months	Grocery shopping : No	Grocery shopping : No access
6d: Aspiring Households	96.5	66.6	84.0	121.0	99.9	83.1
4a: Prospering Younger Families	102.9	100.8	61.6	120.1	103.1	75.3
3c: Accessible Countryside	106.3	102.1	96.2	112.9	98.6	89.8
2b: Settled in the City	88.6	79.0	93.7	111.0	100.4	86.8
4b: Prospering Older Families	101.8	108.7	92.9	108.8	101.3	87.4
6a: Settled Households	98.5	103.7	89.9	105.3	99.8	92.3
2a: Transient Communities	86.7	92.4	45.8	105.3	100.8	85.4
4d: Thriving Suburbs	97.4	90.3	75.9	103.8	104.5	81.7
3a: Village Life	110.6	96.2	96.1	103.7	96.8	99.6
6c: Young Families in Terraced Homes	101.4	93.0	109.0	102.6	98.9	95.4
6b: Least Divergent	99.9	98.5	108.3	102.1	98.6	96.8
7b: Afro-Caribbean Communities	89.3	99.3	103.9	97.6	101.1	92.0
1c: Older Blue Collar	102.4	99.2	87.2	95.6	95.3	107.9
Unclassified	97.0	84.8	94.4	95.6	97.1	102.9
4c: Prospering Semis	100.6	96.2	87.9	94.8	100.0	97.7

1b: Younger Blue Collar	99.7	93.8	83.9	87.6	95.9	111.2
5b: Older Workers	93.0	97.7	103.7	85.4	94.7	115.2
7a: Asian Communities	90.9	87.2	82.3	85.3	102.3	98.6
3b: Agricultural	114.6	113.4	93.5	84.9	99.0	100.1
1a: Terraced Blue Collar	93.1	99.5	113.5	83.2	96.5	110.7
5c: Public Housing	91.6	98.2	104.1	76.9	93.2	119.6
5a: Senior Communities	94.6	96.5	51.0	72.5	86.3	134.5

Source: ONS, 2010

Despite the fact that correlations between ethnicity and shopping attitudes is not as strong as between age, gender and income, it is necessary to comment on this demographic characteristic as the typical UK consumer is more culturally diverse than ever (ONS, 2013). The data from Eurostat in Table 3.2 indicates that there is not much variation in online buying among individuals who originate from the UK and individuals who are foreign born. However, the nationals of 'another EU-Member State' have greater interest in online shopping with 30% purchasing their groceries online in the last three months compared to 18% of individuals who are born in a non-EU country. The multivariate analysis by OAC indicates that Afro-Caribbean and Asian Communities are not enthusiastic online shoppers.

Figure 3.19 shows the distribution of population in the study area by OAC and possible locations of online customers based on this demographic classification.

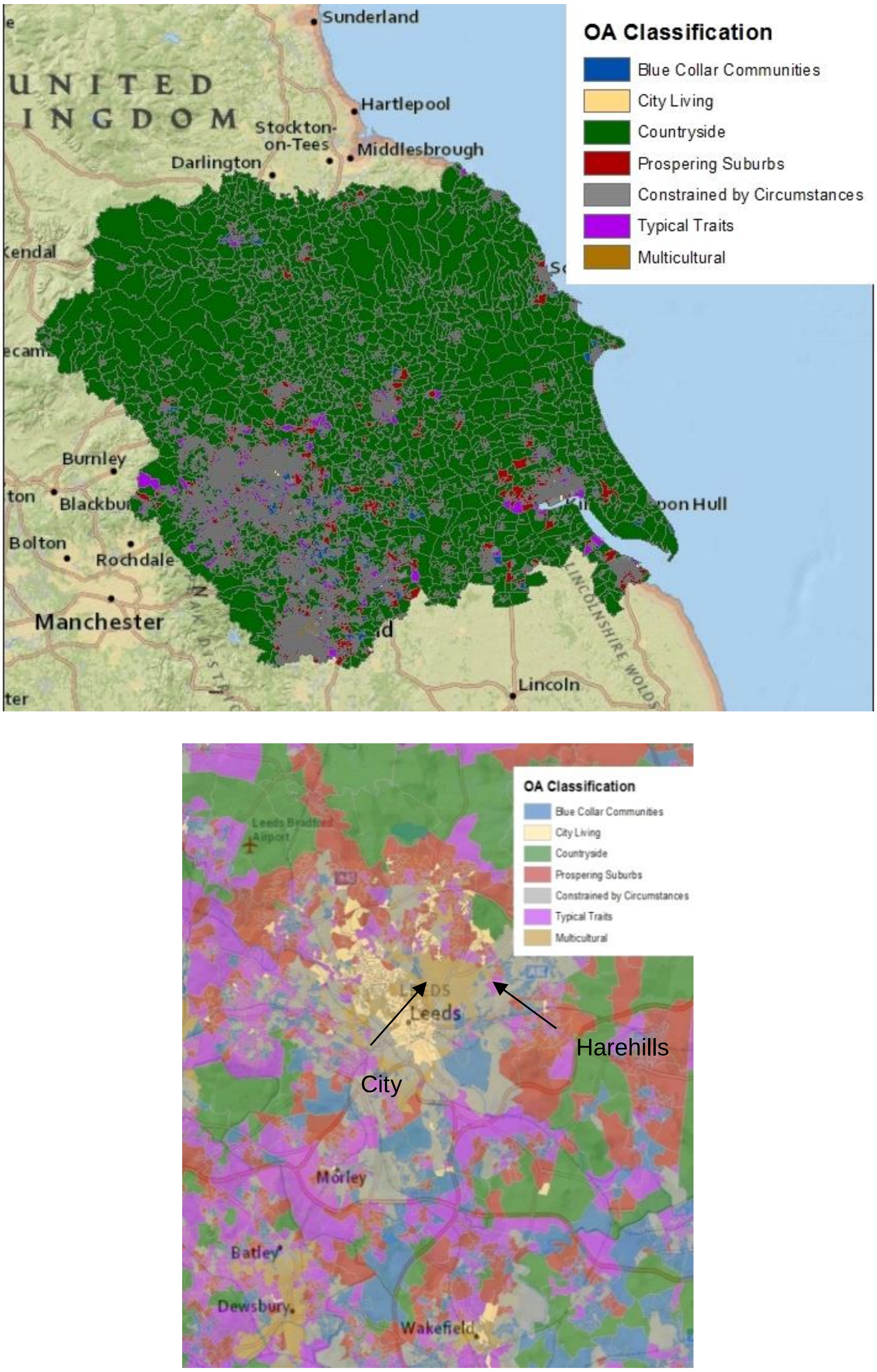


Figure 3.19. Output Area Classification – Yorkshire and Humberside

Due to the nature of the study area, the majority of the customers belong to the Countryside category, which are enthusiastic online grocery shoppers. Therefore, the enlarged map of the city of Leeds provides a better understanding of the likely variation in the distribution of e-grocery customer. It is expected that there will be a higher online expenditure in the city centre locations which have higher concentrations of the City Living category. Higher demand for online grocery shopping is also predicted in the northern and eastern parts of Leeds which are popular among Prospering Suburbs households. Low or very low online expenditure is likely to be in the southern part of the region (areas such as Dewsbury and Batley) and more deprived central areas of Leeds (i.e. Harehills) where there are high concentrations of residents within the Constrained by Circumstances and Multicultural categories.

The analysis of online grocery expenditure within the ACORN demographic classification revealed that the most enthusiastic online shoppers are surprisingly consumers within the Urban Adversity category (Figure 3.20), although, they scored below average in buying groceries online in the last 12 months category. Households within the Rising Prosperity category are also likely to buy their groceries online with an index of 112 compared to the national average figure. The least enthusiastic online grocery shoppers are within Affluent Achievers category which have the lowest index of 70 in weekly online grocery shopping, although, they use online channel occasionally to buy their food, especially bulky and heavy groceries and household goods.

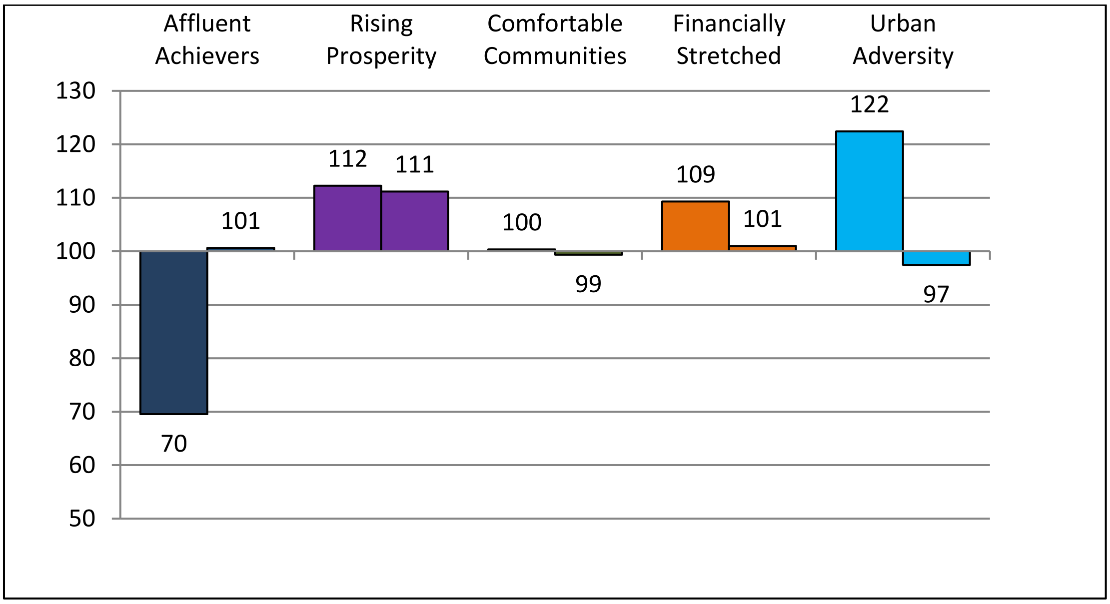


Figure 3.20. ACORN Categories – Online Grocery Shopping – Weekly and in the last 12 months. Source: Acorn, 2013

To create a more accurate profile of the e-grocery customer it is necessary to explore preferences towards online shopping at a more detailed level of resolution. Figure 3.21 shows the distribution of weekly online grocery shopping among ACORN groups. There are 17 groups in total. City Sophisticates are mostly likely to do their weekly grocery shopping using the online channel with an index of 140 compared to the national average consumer and they have the highest penetration within all groups of 15% of total online grocery expenditure (CACI, 2016; Figure 3.20). Their busy lifestyle and low car ownership levels makes the online channel more advantageous compared to the physical in-store channel. Students and young families with modest means are also likely to purchase their groceries online regularly with an index 134 and 124 respectively. According to the latest CACI analysis, poorer pensioners are in fifth place out of 17 in online channel preferences among ACORN group (CACI, 2016), although their relatives or somebody else might be ordering groceries online for delivery at their home address. Nevertheless, it shows that the online grocery market is maturing and the demographic profile of the online shopper is becoming more complex. Moreover, all groups within the Urban Adversity Category scored above average for online grocery shopping. This may sound surprising, although this fact supports the accessibility theory with more customers living in the deprived and less accessible areas employing online channel to fulfil their grocery needs.

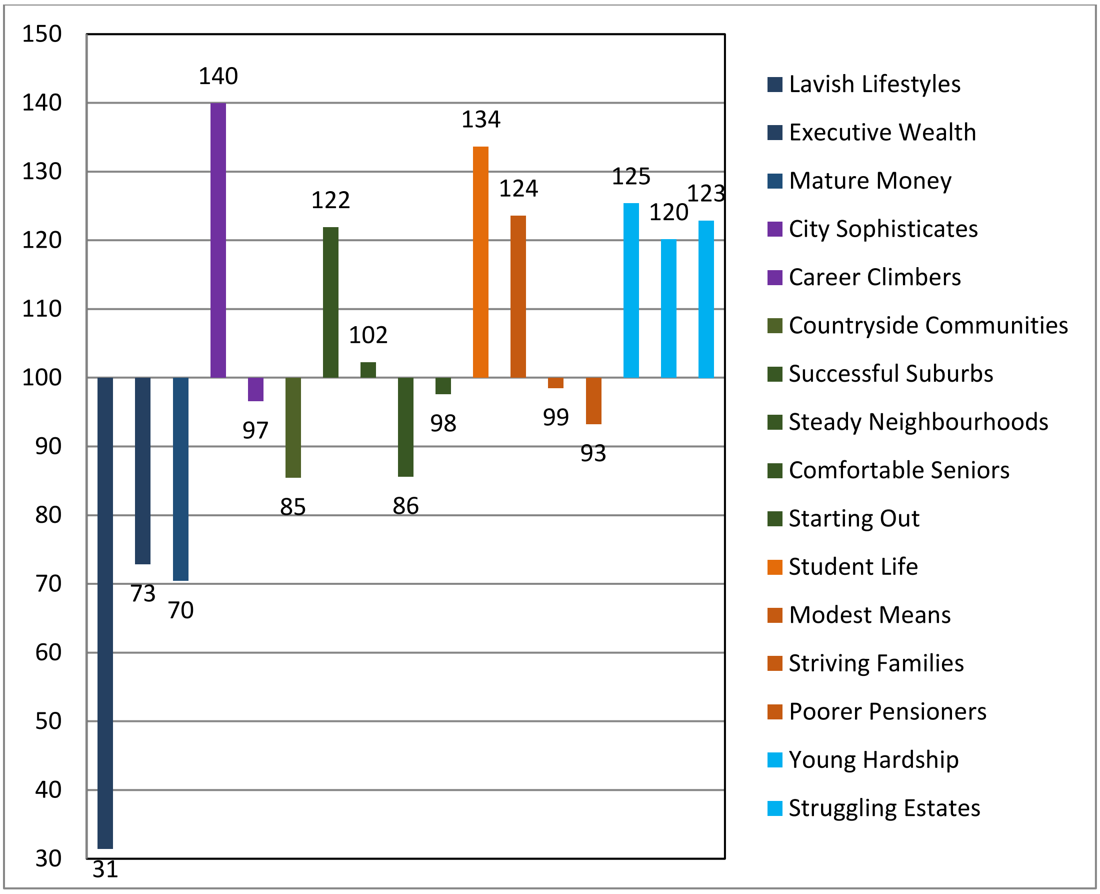


Figure 3.21. ACORN Groups - Online Grocery Shopping – weekly. Source: CACI, 2013

Affluent Achievers are the least likely consumers to buy their grocery online with all three of the top groups scoring below the national average figures in weekly online grocery shopping category. This category contains households with high car ownership levels which makes access to stores easier. Consumers within the Lavish Lifestyles group have with the lowest score of 31 among all ACORN groups but enjoy eating out and consequently cooking much less at home (CACI, 2016).

The distribution of online customers by ACORN categories is presented in Figure 3.22

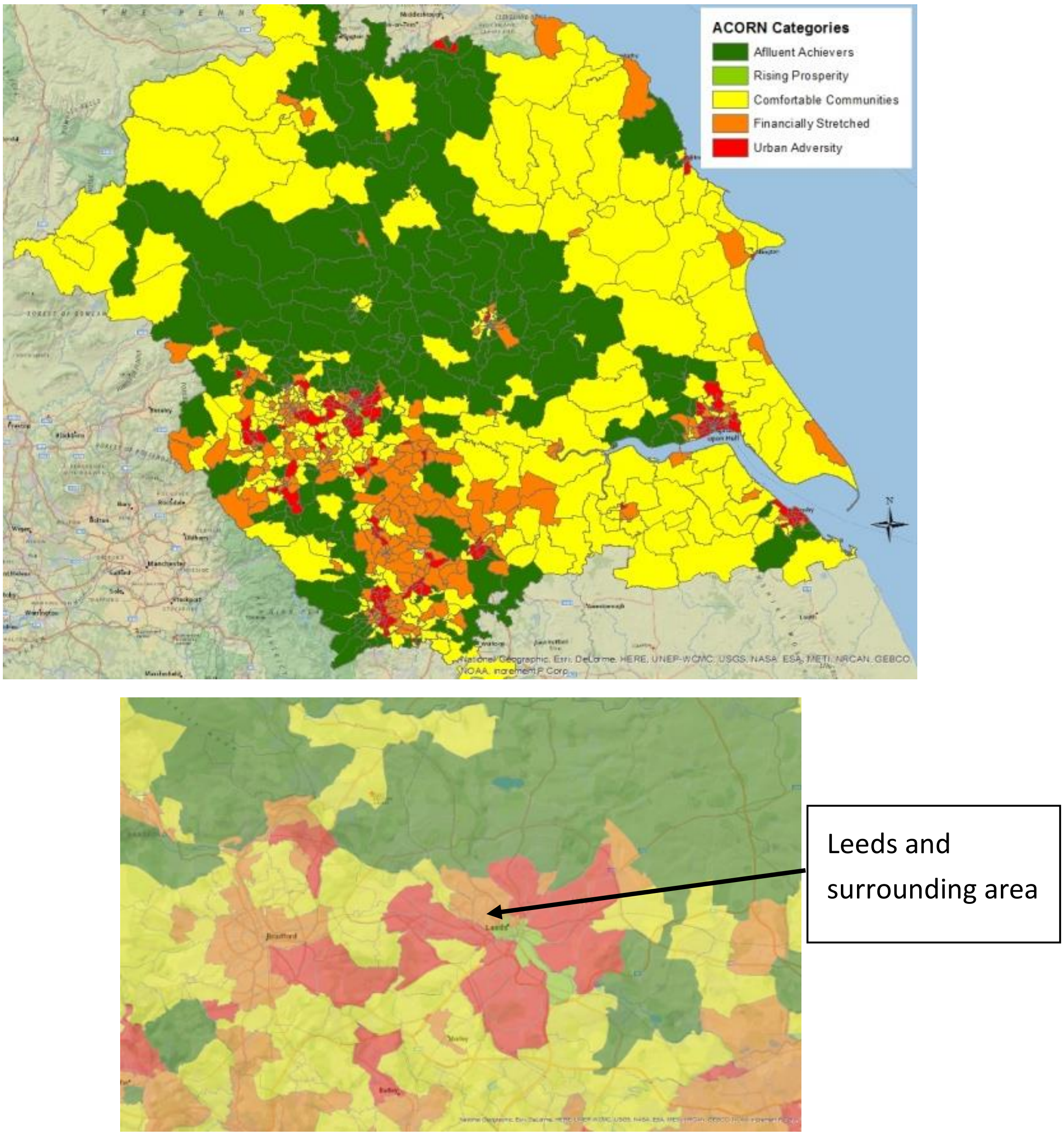


Figure 3.22. Distribution of ACORN Categories – Yorkshire and Humberside

Figure 3.22 shows that the majority of the population in the study area belong to the Affluent Achievers and Comfortable Communities categories which are not enthusiastic online grocery buyers. However, the large proportion of population within Comfortable Communities category belong to Successful Suburbs group which has a higher score of 122 compared to national average in online grocery category. Therefore, the possible locations of online grocery shoppers are in the rural areas of the study area with high concentrations in Successful Suburbs households, urban areas with high proportions of City Sophisticates, Students Life and Urban Adversity customers. The enlarged map of the city of Leeds shows the distribution of online grocery buyers in the urban area. In the heart of the city (more affluent areas with high concentration of City sophisticates) a high online expenditure would be expected (green coloured postal sectors). The less affluent urban areas (red coloured postal sectors) will also generate high online expenditure, although to a lesser extent. Comparing this map with Figure 3.19 there are some similarities and differences of possible locations of online grocery customers. For example, City Living and City Sophisticates (enthusiastic online customers) have similar locations in the urban areas. In contrast, areas with similar locations occupied by Constrained by Circumstances and Urban Adversity are expected to generate low online expenditure.

3.6. Conclusions

This review has revealed that the profile of online grocery shopper is increasingly complex and has changed in a short period of time from largely young professional males living in the city to more mature men or women living in rural locations or urban suburban locations. Convenience is the major factor for choosing an online channel. Therefore, customers with low car ownership and low accessibility to grocery retailers are more likely to benefit from online channels. These categories include students, city professionals and families with children. Generally, there are three types of online grocery buyers. The first type is a young professional single or couple living in affluent urban areas with a high income. The second type is less affluent households of all ages on low incomes living in urban areas with limited access to physical stores. The third type is middle class (skewed towards more affluent social classes) families with young children living in the rural or suburban areas. To be successful in the online environment grocery retailers need to consider strategies which will target particular types of online customer. For example, students and pensioners households (type one) are frequent online customers but they do not spend a large amount per single transaction. They tend to buy regularly but spend less. By offering special promotions (for example, reduced delivery charges or complete free delivery) to loyal customers they will minimise fulfilment costs as this particular demographic groups prefer less affluent supermarkets and tend to buy from the same retailer.

It has been established that demographics play an important role in establishing of possible locations of online customers, although, of course, other factors are also significant in decision making process of online purchasing such as attractiveness of the brand and the cost. The demographic patterns are complex and geodemographics is one technique which enables representation of this complexity. The existing demographic classifications (i.e ACORN and OAC) have advantages and disadvantages in their application. This research uses both of these techniques to identify profile of online grocery shopper. Moreover, the appearance of multifaceted social trends (e-shopping) required development of bespoke demographic classifications applicable to small area of geography (e.g. postal sector) which can be aggregated to a larger scale (e.g. county or region). Generally, the choice of demographic typology is not significant as all of the recently developed classifications are based on multivariable demographic characteristics and reflect complexity of customer's behaviour. Further chapters will explore the demographic profile of online customers based on actual online expenditure data provided by client. This chapter, and the literature review in Chapter 2, has provided an overview of the data based on various surveys and estimated data. The question is, will demographic analysis of online customer based on actual online expenditure match these demographic profiles? This will be discussed in Chapter 6 and will provide the first academic study to attempt to answer this question.

Chapter 4: Review of site location research techniques

4.1. Introduction

This chapter will review the existing models typically used for retail site location and justify the chosen model for the purpose of this research. First, section 4.2 provides an overview of the development of site location practices in the UK, concentrating mostly on recent contributions. Section 4.3 provides a comparative analysis of the existing site location techniques and a justification for the applied Spatial Interaction Model (SIM) technique used in this thesis. Section 4.4 outlines the implications for these types of models of the major transformations in the retail environment given changing consumer behaviour, with an emphasis on new consumer movements and the effects of new technologies on the space and time variables (especially of course the growth of online retailing). Finally, section 4.5 describes the components of the SIM and its benefits and limitations.

4.2. Development of site location techniques in the UK

It was not until the mid to late 20th century when the first retail location techniques started to emerge. In the 1970s retailers were still mainly relying on 'gut feel' or the intuition of experienced managers to determine the location of their next store (Reynolds and Wood, 2010). That said, the first attempt to theorise site location models was made in the 1930s by an American industry worker, David Applebaum, Chief of the Market Research Department of the Kroger Grocery and Baking Company, with his work on analogue models and customer spotting techniques. Since these humble origins, site location techniques have become more analytical and more reliant on quantitative analysis. Indisputably, they have improved the productivity and the effectiveness of the retail industry. However, some concerns started to be raised regarding excessive dependence on technological methods with retailer experience becoming increasingly underestimated and overlooked. Many researchers agreed that a combination of quantitative and softer qualitative methods is required to produce sophisticated estimation models, although this creates a difficult challenge for site location analysts (Clarke et al, 2003). Moreover, some researchers urged retailers not to concentrate their efforts purely on the maximisation of sales but also to consider the retail operation as a whole unit with finance, logistics and other departments equally important for affecting business profitability (Walters, 1974).

The rapid and widespread development of large supermarkets and superstores, offering an extensive spectrum of products and attracting customers from wider geographical areas, created new challenges for planners to forecast sales, with some researchers raising concerns regarding retailers making huge investments on large stores without detailed methodological considerations for new site locations. For example, Simmons (1978) argued that retailers may endure financial losses due to the lack of knowledge of the trading potential of any new store. To eliminate the risk of failure of a new store, he suggested performing post opening analysis to generate better future sales forecasts and as a result improve the effectiveness of the new store planning process. New techniques started to emerge to identify potential locations for new stores, so called 'search' and 'viability' techniques which help to predict sales of the store based on a particular site (Reynolds and Wood, 2010). These techniques became very relevant since the changes around planning regulation and growth of out of town stores.

In the late 1980s, the beginning of the fast emerging technological advances, many scholars observed that with the simple spreadsheet retailers can employ a powerful tool in site retail analysis. As technologies became more accessible and affordable so did the productivity and applicability of new site location models. By late 1990s many major retailers possessed a combination of methodologies for sophisticated site location analysis including Geographical Information System (GIS). Despite the popularity and reliance of planners on the complex technologically-driven modelling techniques, some researchers commented on the importance of network planning and site analysis at the micro level (Alexander et al, 2008). Moreover, understanding shopper's motivations also determines the effectiveness of site location (Davies and Clarke, 1994). For example, grocery shopping will create different patterns compared to non-food shopping trips. Furthermore, the size of the goods (heavy and bulky versus compact and light) will affect the profitability of the store. Wood and Brown (2007) noted that existing conventional models are not effective at the micro level especially in forecasting convenience store sales. They stated that a combination of practical knowledge and office based research would be necessary to design more accurate models to predict store sales. Other researchers have come to similar conclusions with the need for a balance of subjective (intuitive) and objective (data collection) approaches to develop sophisticated site location techniques (Simkin et al 1985; Hernandez and Bennison, 2000).

Reynolds and Wood (2010), in their analysis of existing site location methodologies, established that the size of the planning location team is correlated to the size of the business. For example, typically the team with 10 or more planners is managing a retail business with more than 500 stores (with the tendency of larger businesses to rely on their own in-house team of planners) whereas smaller retail organisations prefer to outsource the design of forecasting models to external agencies. Moreover, their survey of more than 102 individual businesses revealed that retail location teams are more focused on site assessment for new individual stores and do not have a great deal of involvement in the more strategic divisional side of site location assessment. Generally, the site location teams deal with only new and relocated stores and do not make decisions regarding the particulars of the store, e.g. staffing level and store format. Furthermore, the analysis of the use of different location techniques (from 1998 to 2010) revealed that some techniques seem to have gone out of fashion with some methods becoming increasingly more accepted. For example, nowadays analogue models, cluster analysis and checklist approaches are less employed by retailers compared to more popular gravity modelling and multiple regression analysis. This partially reflects the concentration on the development of new stores in the last ten years. In 2010, 82% of the respondents (retailers) reported using GIS compared with 53% in 1998 (Reynolds and Wood, 2010). However, the use of new technologies is not generally available through the entire organisation. The researchers suggested that in order to increase the overall effectiveness of the business a greater involvement of retail location teams is required, as they collect an extensive data base which can have a greater use in all aspects of business operation, not only limited to the site location processes. In their research Reynolds and Wood concluded that grocery retailers have the most sophisticated spatial analysis techniques (which are usually implemented by an in-house planning team) managing a likely portfolio of more than 500 stores. Moreover, grocery retailers have access to loyalty card schemes and Electronic Point of Sale (EPOS) data which provides extensive data on customer's buying behaviour and their geodemographic attributes (Birkin and Clarke, 2009).

Changing consumer behaviour and complex lifestyles create new challenges for site location planners in terms of demand estimation and, consequently, choice of store format and product range. For example, Newing et al (2013) adapted SIMs to include seasonal variations of tourist demand in parallel with the more static residential demand for grocery retailers in Cornwall with a high accuracy of sales estimation. Hood (2016) has designed new modelling techniques to estimate very dynamic consumer behaviour for the convenience market with an application based on consumer cluster analysis and SIMs. Emerging social trends and changes in consumer attitudes will create further challenges for planners with the need to develop more sophisticated models which will employ technological advances, data collection, and analytical skills in parallel with common sense, experience and intuitive approach. These models will be reviewed later in the chapter.

4.3. Review of site location research techniques

There are various site location models which are widely recognised and accepted by geographers and retail planners worldwide. These models use either a deductive or inductive approach and can have simple or complex methodologies (Birkin et al, 2002). For example, agent-based modelling method is a deductive approach as it focuses on individual's behaviour, identifies common characteristics, aggregates the data and creates a model which will predict behaviour of a group of individuals in a particular situation. In other words, is a bottom to top approach as opposed to SIMS, which follow an inductive approach (moving more from top to bottom).

The simplest technique for site location is analogue modelling. This approach is based on comparing new or existing stores to a similar store in the UK in relation to store's characteristics, e.g. size, location, catchment, population, etc. Analysists can make predictions in two ways. First, the proposed store's sales are evaluated against the analogue stores within the corporate group. This can be achieved by a company reviewing its sales figures and using a ranking system according to total sales or sales per sq ft. Also a company can rank revenues according to store formats (supermarkets, convenience stores) or by location types (rural via urban sites) (Birkin et al, 2012). It allows retailers to evaluate various store types across the corporate chain and apply the format which will be more profitable for the store in the particular area. For example, locations with no competition and an extensive target market will be appropriate for superstores with a large selection of products.

The second approach of this technique relates to planning opportunities. If a retailer sees an opportunity to acquire a new site it may apply an analogue model to find an existing store with similar geodemographics to the proposed site. Based on the performance of this existing store a retailer will make a decision on the new site acquisition. The drawback of this technique is the difficulty of finding existing sites with very similar characteristics to the proposed store and then the question of whether these characteristics can be applied to the new location successfully. In such competitive and varied environments, grocery retail planners face a serious challenge to find two location sites with identical or very similar characteristics. The success of analogue modelling depends on the experience of site location teams and the availability of the site locations themselves with analogue characteristics and similar geodemographic profile of the stores' catchment areas. Moreover, the existing comparable store may be under- or over-achieving, which will not therefore create problems for producing accurate performance statistics for a new store (Birkin et al, 2002).

Geodemographic classification is a similar technique to analogue modelling, but is based more on the categorization of populations within the catchment area of a new or existing store. The particular area unit, (e.g. postal sector) will have a certain population profile, e.g. "young professional". Locations with similar profiles can be identified and new stores can be expected to achieve the same type of performance as the analogue stores within the same geodemographic profile catchment area.

Ratings models are more appropriate in complex and volatile markets with a wide range of customers and various distribution channels (Birkin et al 2002). Site location for ATMs is a good example for an application of this technique (ibid). The decision around finding good site locations will depend upon the overall rating score based on key elements, e.g. market size, demographic mix and supply ration, which are individually rated.

Statistical modelling is a more advanced technique in comparison to analogue modelling and involves the statistical methodologies of correlation and regression. Regression analysis is based on finding correlations between a dependent variable (i.e. store revenue) and set of independent variables (e.g. distance to store, retail brand, size of store). Linear regression is perhaps the simplest method to establish the relationship between two variables, e.g. store revenue and various attributes (store size, competition, market size). The reliability of the results is tested against the R-square value, which should be closer to 1 to validate a clear correlation between two variables. Birkin et al (2002) point that the r-square value and other statistical measures are insightful but they proposed to employ "the average error of the predictions". This value will indicate the error margin of the model, e.g. 10%. The stakeholders will decide for themselves if this rate is significant or not for their purposes.

The basic regression model can be described as follows:

$$Y_i = a + b_1 X_{1i} + b_2 X_{2i} + b_3 X_{3i} + \ldots + b_m X_{mi} \quad (4.1)$$

Where Y_i is the grocery store *(i)* sales (the dependent variable) and X_{mi} are independent variables (e.g. customer demographic characteristics). b_m and α are regression coefficients and the intercept parameter respectively.

Multiple regression techniques allow the testing of the effect of various attributes on the variable to be predicted, e.g. store revenue. The sequence is formed in order of the significance of the attributes, starting with the most influential attribute. In network planning regression is widely used. For example, Duggal (2007) explores factors which have the most effect on the US fast food restaurant revenues (his examples include McDonalds and Burger King). The multiple regression analysis established that ethnic population and household income are the most important variables for both companies, although, the analysis also established that McDonald's customers are willing to travel longer distances compared to Burger King's customers.

However, this technique has a few setbacks. First, the regression modelling for physical stores does not consider the complete impact of the competition (Birkin et al, 2002). Secondly, it is based on the assumption of similarity of store samples, which will be difficult to achieve (as with the analogue approach). However, the major disadvantage of this method is that it does not consider customer flows in detail, i.e. it does not allow the estimation of revenues generated by spatial interactions and movements of customers between home, workplace and retail units (ibid). Statistical techniques are often part of more advanced GIS methodologies and used to test various combinations of GIS variables for good correlations between them, e.g. store revenue and an area's geodemographic characteristics (age, social class, family composition). However, overall, in the case of high correlations between existing data of sales and geodemographics, the revenue of new stores can be estimated with high level of accuracy.

There are other multivariate statistical methods which can be applied in network and store planning. An interesting example is the study by Simkin (1985) who applies factors of competition, site areas, accessibility, store and catchment area characteristics to create revenue estimation for different retailers. His model produced high correlation (r-squared) of 0.81 and above for major retailers in electronics, fast food industry and dry cleaning.

Geographical information systems (GIS) is designed to capture, analyse and present all types of geographical data (Birkin et al, 2012). GIS allows the user to identify potential target markets for existing and new sites by applying the 'buffer and overlay' technique. This method has two stages. First, it determines how far consumers are willing to travel to the store which can be done through travel time or straight line distance. Secondly, a buffer is drawn around the store based on the travel distance, thus defining the catchment area in all directions around the store. The store revenue is then estimated based on the amount of residents and the residents' income within the buffer. The drawback of this method that it assumes that customers living at the edge, in the middle, and close to the store are all as likely to travel to this particular store. Store revenue prediction becomes even more difficult when the competitors stores within the buffer catchment area are also considered.

Despite its popularity GIS has a few disadvantages. It depends on the quality and validity of the geographical data. For example, there is a difficulty in identifying the catchment area for a new store and how to take account of the rival stores in the area (Birkin et al, 2002). This method usually operates a "fair share" system, where the market share is distributed equally between the competing stores. In reality that is unlikely to be the case as store attractiveness will vary by size, brand, accessibility etc.

The spatial interaction model (SIM) connects two attributes - demand and supply for retail site selection. Moreover, it takes into account a third factor – accessibility to the store. The more detailed analysis of SIM will be provided in the next section as this is the model used in the rest of the thesis. In general, the model is based on the assumption that patterns of customers flows depend on the attractiveness of the store and its degree of accessibility.

An alternative approach would be to use micro models, i.e. agent-based modelling (ABM), which is the computational study of social agents as individuals with complex interaction patterns (Janssen, 2005). ABM uses a deductive approach, i.e. information gathered about individual behaviour is aggregated into macro data. By using an ABM a researcher can identify agents' common behavioural patterns depending on their attributes and their effect on the macro level. The major advantage of ABM is the consideration of the possibility of complex social interactions. Despite its obvious benefits ABM has drawbacks. For example, ABM studies individual behaviour which potentially can be irrational and subjective. These factors make it difficult to quantify and calibrate the data, especially at the aggregate level (Bonabeau, 2013). Furthermore, the process of describing and simulating an individual agent's behaviour (and then aggregating it to a macro level) is a very time-consuming process which makes it difficult to apply to a large modelling system (ibid). However, an interesting future project would be to fuse SIM with agent-based models to create a set of individual SIMs (see also the conclusions in Chapter 9)

4.4. Changes in location planning techniques

Chapter 2 demonstrated that the modern consumer tends to be a highly informed and sophisticated individual with demanding needs who is becoming increasingly time conscious leading an increasingly complex lifestyle. This creates difficulties for retailers and marketers to predict modern consumer movements. The conventional techniques to calculate demand using customers' demographic characteristics and their locations is becoming outdated and planners require new more advanced methods to incorporate contemporary customers' accessibility to various modes of shopping

(Cliquet, 2006). The study of consumer behaviour is a very popular subject among researchers from various disciplines, especially marketing, with half of all marketing research undertaken using behavioural studies. What is lacking in the literature is the combined study of space or geographical location and new behaviours, i.e. *new spatial behaviour*. In future retail planning research this new consumer behaviour requires application of two principles - retail location and marketing management which includes a greater reference to the marketing mix (product, price and promotion). Most existing spatial behaviour models apply a static approach (with fixed locations) and do not consider new consumer movements, especially around concepts such as e-commerce.

That said, despite the difficulties facing retail planners, there are three fundamental principles of store attraction which stay unaffected in consumer behaviour - variety of goods or services, convenience and accessibility. In the traditional SIMs (see section 4.4) the variety of goods and services is represented by the size of the retail unit with the larger format stores considered to be more attractive with a wider choice of products. Copeland (1923) identified the following three product categories which influence consumer selection of a store, and this classification is still applied by modern retailers:

- Convenience goods which are regularly purchased locally
- Shopping goods which require a more detailed information search
- Speciality goods which can only be purchased at a particular store selling branded goods

Key questions include how have new innovations and technologies changed location modelling techniques? The emergence of the Internet has introduced a new dimension to the concepts of measuring distance, accessibility and product selection. The Internet may have eliminated space and time constraints to some extent (Farag, 2006), but at the same time customers and businesses remain dependant on their physical location and difficulties it may present, e.g. delivery of goods to remote areas (Cliquet 2006).

It may appear that the development of new technologies in the last 50 years (mobile phone, fax and Internet) have eased consumers' lives and freed them from making physical trips to shops. But research shows that new technological advances have not actually reduced the number of shopping trips, but in fact have had the opposite effect by creating more trips (Allemand, 2001). The multi-channel (or omni-channel phenomenon) introduced in previous chapters has been studied by retail scholars and attempts have been made to conceptualise this new customer behaviour. Recent studies have identified that the choice of shopping channel is influenced by its perceived safety, value for money, accessibility, search effort and delivery time (Gupta et al., 2004; Gong and Maddox, 2011). Javadi et al (2012) identified that online customer behaviour is related to demographic characteristics, online expertise, channel's convenience and shopping rationale (see Chapter 3). Zilmans (2010) also confirmed that the preference of the shopping mode is associated with characteristics of the consumers, products, shopping channels and the retailers. Conclusively, the choice of channel ultimately depends on two factors: customer traits (demand) and stores characteristics (supply). SIMs in theory do consider these factors and consequently should still be a potentially useful modelling technique to take account of these new developments in consumer behaviour. However, there will need to be some interesting modifications to model structure – a major aim of this thesis is to offer new directions in this area of research.

4.5. Spatial Interaction Model (SIM)

In the last twenty years the SIM has become a key site location modelling technique for retailers, offering a more sophisticated methodology to estimate store revenues in a very competitive, arguably saturated environment. The SIM is based on two fundamental assumptions: places with large populations are inclined to generate more activities and more remote places generate less interaction. SIM is a type of gravitational model and based on the study of retail interactions which are determined by two factors – supply and demand. Supply comprises of the products within a set of stores and demand is determined by consumer behaviour. Gravitation models rely on two facts – mass and distance (Cliquet, 2006). The early attraction models were based on the Newtonian law of gravitation with its deterministic approach. For example, Reilly (1929) introduced the basis of the theory of spatial interaction stating that consumers will be attracted to a particular retail store based on its attractiveness against the distance travelled to that store. This 'law' does not consider overlapping retail trade areas and, therefore, is not very suitable in urban areas for retail trade estimation. The later models have attempted to overcome these obstacles by introducing amore probabilistic approach. Huff (1963) designed a gravity model which included both essential parts – distance (accessibility time) and mass (size of the store or floorspace). His model included variables which also measure the competition and attractiveness of different stores based on their size and product range.

Despite its originality and its probabilistic approach the Huff model had limitations in terms of homogeneity of the consumers and the stores, with no differentiation of consumers' geodemographic attributes and store's brand attractiveness. The later SIMs models became more sophisticated in capturing consumer attributes, and were influenced by the work of Alan Wilson in the 1970s. He replaced earlier Newtonian analogy models by building a suite of models from first principles using entropy maximisation. Thomas and Hugget (1980) stated that entropy maximisation has a behavioural significance:

"*when we construct entropy maximisation models it is assumed that we will never find out which route each of the individuals actually assigned themselves. Given this assumption, the entropy maximisation criterion has a behavioural meaning because we select the solution that maximises the individual's freedom to choose between available trips. For this reason the entropy maximisation solution is said to be the most likely trip matrix*"

(p.156)

Moreover, entropy maximisation models introduced constraints and balancing factors which overcome criticism of early gravity models as being too aggregate with inadequate forecasting ability. Wilson (1971) introduced 'the family of spatial interaction models' which are differentiated by the constraints which are placed on each member of the family. He proposed four scenarios based on different combinations of $O_i$and D_j :

1. The unconstrained case where neither O_i and D_j are known;
2. The production constrained case where O_i is known and D_j is unknown;
3. The attraction constrained case whereO_i is unknown and $D_j$is known;
4. The production-attraction constrained case where both O_i and D_j are known.

For retail applications the production constrained model is the most appropriate where O_i can be defined as the expenditure available in origin zone *i* and the mass of the destination D_j can be replaced by store attractiveness (floorspace) W_j in order to estimate the revenue D_j the flows from an origin zone *i* is constrained to the available expenditure in that zone (demand), whereas the flows to the destination zone *j* are unconstrained.

The classic production constrained entropy model can be is represented as follows:

$$S_{ij} = A_i O_i W_j exp^{(-\beta C_{ij})} \quad (4.3)$$

Where:

S_{ij} is the flow of people or money from residential area *i* to retail unit *j;*

O_i is a measure of the available demand (grocery expenditure in this case);

W_j is an attractiveness factor for retail unit *j (*i.e. size)

C_{ij} is a function representing the cost of interaction between demand zone *i* and store *j,* most commonly in the form of straight line distance between the two

A_i is a balancing factor ensuring that all demand is allocated between the available grocery stores and is calculated as:

$$A_j = \frac{1}{\Sigma W_{j\ exp}{}^{-\beta C_{ij}}} \quad (4.4)$$

$exp^{(-\beta C_{ij})}$is the form of the distance deterrence factor most widely used where C_{ij}(the distance between origin zone *i* and destination zone *j)* is influenced by an additional parameter $-\beta$. (Birkin et al, 2002; Birkin and Clarke, 1991; Wilson 1971)

The distance deterrence parameter β measures the customer ability and desire to travel to the store. Generally, β will be higher for low cost convenience products (groceries, newspapers) and lower for more expensive goods (cars and furniture) as customers are willing to travel longer distances for those products and the obstacle of distance is less important. The β value also depends on the coordinate system being used. The distance becomes very large between two zones with the application of an exponential function (as seen in equitation 4.3) and the six figure OS co-ordinates (Clarke and Birkin, 2016). In this case the β value will also need to be set according to actual distances travelled.

Thus, in the retail model, the demand O_i from geographical zone *i* is distributed across available retail units j based on their accessibility and respective attractiveness factorsW_j. Demand is normally expressed as household expenditure in that neighbourhood derived from national statistical data (Birkin et al, 2010b). The supply side of the model is represented by the attractiveness of the stores, with the available floorspace being the main attractiveness factor used in the literature. Generally, larger stores have higher attractiveness scores compared to smaller retail units. However, other store attributes can also be important, e.g. available parking spaces, range of products, opening hours and store fascia. Moreover, the location of smaller stores in a well-established centre may be more attractive to consumers compared to larger stores in a standalone unit (Birkin et al, 2010b). The other very important factor in consumer purchasing decision making process is brand attractiveness. For example, consumers within the higher social class group may prefer to travel longer distances to Sainsbury's to do their grocery shopping despite a close proximity of an equal size ASDA supermarket. Consequently, size alone is not the decisive factor of store attractiveness which is usually a combination of factors including multiple store attributes related to the socio demographic characteristics of customers. To calculate the overall attractiveness of a store the scorecard technique may be appropriate which includes multiple characteristics of the store (Birkin et al, 2010b). A more comprehensive analysis of the demand and supply side of the model will be provided in Chapter 7. The model works on the assumption that consumer choice of the equally accessible stores will depend on the store attractiveness. However, these preferences are not deterministic as consumer will not necessary choose the most attractive store between the equally accessible stores. Consequently, the model has ability to reflect more complex customer behaviour.

The cost of distance can be measured as straight line distance which does not reflect reality with complex road networks and traffic congestion. The more sophisticated technique is to produce travel time matrices based on average speeds for these networks with consideration of likely obstacles (rivers and motorways).

To reflect more complex consumer behaviour and attributes, the model presented in equitation 4.3 can be disaggregated by different household types (m) and store attractiveness (α) as follows (Clarke and Birkin, 2016):

$$S_{ij}^{m} = O_i^m A_i^m W_j^{\alpha^m} \exp(-\beta^m d_{ij}) \qquad (4.5)$$

where:

S_{ij}^{m} is the flow of people or expenditure disaggregated by household type *m*

O_i^m is the demand of household type *m* in residence zone *i*

A_i^m is a balancing factor which is calculated as:

$$A_i^m = \frac{1}{\sum_j W_j^{\alpha^m} \exp(-\beta^m d_{ij})} \qquad (4.6)$$

W_j is store attractiveness of destination *j*

α^m is a parameter of store attractiveness by household type *m*

d_{ij} is the distance between origin *i* and destination *j*

β^m is the distance decay parameter for household type *m*

Despite its advantages and sophistications, SIM has its limitations and potential inaccuracy in three areas – applied data, geographical zones and level of disaggregation. The quality and representativeness of data is a very important issue which could undermine the subsequent analysis. To overcome this problem data from various sources should be compared and contrasted to highlight any inconsistencies. Moreover, statistical methodologies may be applied to discover the significance of the data with the use of p-value, for example. The predictions of the spatial interactions is also dependant on the choice of geographical demand zones due to their non-static nature (Fotheringham and Wong, 1991). The level of disaggregation is related to the requirement of additional data at the micro level, which becomes problematic in assessing the importance of data and increases probability of error (Openshaw, 1976). Data errors are likely to emerge in the following areas. First, demand estimation is usually based on sample surveys which can use different methodologies and do not consider the flexibility of customer movements. Therefore, demand is calculated based on static data related to the residential addresses of the consumers. Secondly, errors in the supply data may arise through the narrow definition of the attractiveness of the store. As mentioned previously, size of the store is not a deterministic factor for individual store level site attractiveness. There is no universal rule in the selection of variables. Each individual SIM should be adjusted to each particular task based on the area's specific attributes, e.g. location, customers' demographic characteristics and lifestyles.

To identify and eliminate some of these errors, analysts employ a calibration process which involves choosing the best parameters to obtain the closest match between estimated and actual (or known) datasets. Calibration uses statistical methods to determine values which offer the closest match to observed interaction patterns. For example, different patterns of customer flows can be achieved by changing β values. The well calibrated model can replicate customer flows from demand areas to supply locations very accurately. This is very important for the retailers who will want high level of accuracy for revenue predictions of new stores.

As noted above, the calibration process works on minimising the difference between actual or observed data and predicted data produced by the model which can be presented as follows:

$$\text{Minimise } S = \Sigma_{ij}[S_{ij}(\text{obs}) - S_{ij}(\text{pred})]^2 \qquad (4.7)$$

To determine the gap between the predicted and observed data many goodness of fit statistics can be applied with R^2 being the most commonly accepted method. This can be represented as follows:

$$\boldsymbol{R^2} = \left[\frac{\Sigma_i \Sigma_j (\boldsymbol{S_{ij}} - \bar{\boldsymbol{S}}_{\boldsymbol{o}})(\hat{\boldsymbol{S}}_{\boldsymbol{ij}} - \bar{\boldsymbol{S}}_{\boldsymbol{m}})}{\left[\Sigma_i \Sigma_j (\boldsymbol{S_{ij}} - \bar{\boldsymbol{S}}_{\boldsymbol{o}})^2 * \Sigma_i \Sigma_j (\hat{\boldsymbol{S}}_{\boldsymbol{ij}} - \bar{\boldsymbol{S}}_{\boldsymbol{m}})^2\right]^{1/2}}\right]^{\boldsymbol{2}} \qquad (4.8)$$

Where $\bar{S}_o$ is the mean of the S_{ij}s (estimated data) and $\bar{S}_m$ is the mean of the $\hat{S}_{ij}$s (observed data). R^2 parameter has values between zero and one. The value of R^2 which is closer to one has the closest match to the actual value. The value of R^2 which equals one means a 100% correspondence with the observed value. A zero value reflects no correlation to the actual data.

The SIMs in this research were calibrated against actual data derived from the nectar loyalty card scheme used by the supermarket chain. The more detailed explanation of SIM calibration will be explored in Chapter 7.

The SIM is the most applied modelling technique in the retail sector with over 60% of the site location planners making use of it (Reynolds and Wood, 2010). Moreover, this technique brings quantifiable profits to businesses with a high level of the sales estimations and return on investment compared to other site location techniques (Birkin, 2010). The SIM will be applied in this research but will be modified to account for Internet usage. This will being some interesting challenges and will be the focus of Chapter 8.

Chapter 5: Study area and data sources

5.1. Introduction

This chapter provides an overview of the study area of Yorkshire and Humberside and the data sources used in the rest of the thesis. Firstly, section 5.2 describes the features of the study area in terms of population density, economy and demography. Section 5.3 provides an overview of two major contributors to this research in terms of data sources – CACI (information technology company) and the UK leading supermarket chain. Section 5.3.1 introduces CACI and its research methodology and presents data from this reputable source. Section 5.3.2 provides an overview of the loyalty card scheme and its impact on the retail industry; introduces supermarket chain and its place in the UK grocery market and presents a description of the data received from the site location team.

5.2. Study Area

The study area is the Yorkshire and Humberside region, which is fifth largest region in England with 15400 sq. Km and covers just over 6% of the entire area of the UK. According to ONS (2013a), in 2012 the region had a population of 5.3 million or 8% of the total UK population. The population density in the same year was 344 people per sq. Km (compared to 401 for England as a whole), although it varies widely from 36 people per sq. km in Ryedale, North Yorkshire to 3,700 people per sq. km in Kingston upon Hull (see Figure 5.1). The majority of the population (82%) live in the urban areas (ONS, 2013b). The rural areas are situated in the north and east of the region, with urban areas generally in the south and in the west. Two National Parks, the North York Moors and the Yorkshire Dales, make the region distinctive from other UK regions as they cover 20% of the region, larger than the area of any other National Parks in England. There are seven cities in the region - Leeds, Sheffield, Bradford, Kingston-upon-Hull, York, Ripon and Wakefield, with Leeds being the 3rd largest city in the UK (ibid).

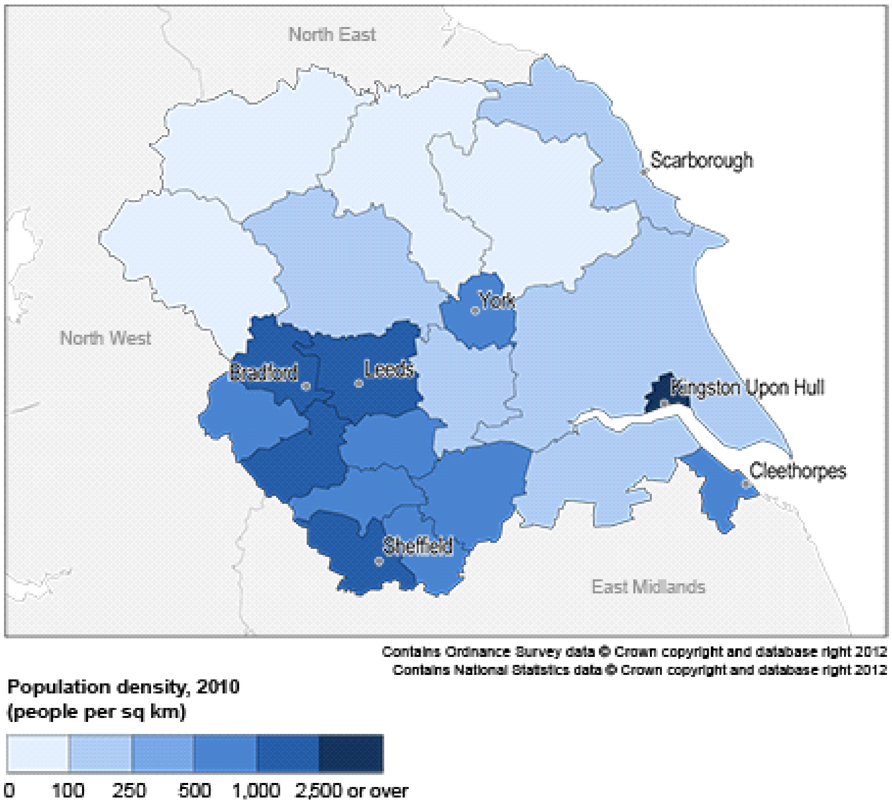


Figure 5.1. Population density: by local or unitary authority, 2010. Source: ONS, 2013a

The median age of the region was 39 in 2012, which is the same as in England as a whole, and ranges from 34 in Bradford to 47 in Craven (ONS, 2013a). The majority of the population belongs to the "white" ethnicity group (89%), which is comparable with an indicator of 86% for the whole of England. That said, the region has the highest proportion of Asian/Asian British with Pakistani residents make up 20% of the population in Bradford. The table below provides a summary of the region's geodemographic characteristics.

Table 5.1. Region and county profile, 2012.

Ethnic group	Thousands	%	Males	%	Females	%	Children under 16	%	Persons aged 16-64	%	Persons aged 65 and over	**%**
White: English/Welsh/Scottish/Northern Irish/British	4,531	86										
99Asian/Asian British: Pakistani	226	4										
White: Other White	130	3										
Asian/Asian British: Indian	69	1	2,602	49	2,686	51	999	19	3,408	65	882	17
Black/African/Caribbean/Black British: African	46	1										
Asian/Asian British: Other Asian	40	1										
Mixed/multiple ethnic group: White and Black Caribbean	33	1										
Asian/Asian British: Chinese	28	1										
White: Irish	26	1										
Mixed/multiple ethnic group: White and Asian	26	1										
Total population	**5,284**		2,737	52	2,547	48	1,006	19	1,399	26	937	**18**

Source: ONS, 2013a

Overall the study area offers contrasting geodemographic characteristics with comparable proportions. For example, the population is equally distributed among rural and urban areas, with half of the population living in the seven cities. The population is equally distributed between the genders and youngest population group (children under 16) and persons over 65 have analogous proportion of 19% and 17% respectively in the total population.

Moreover, the region's attributes corresponds with the UK and England's average indicators. The region is one of the more deprived regions in England with Kingston-Upon-Hull area ranked 11th among the most deprived districts (Communities and Local government, 2007). The rate of crime committed against households in 2013 was 216 per 1000 households which is the same rate as in overall England region of 217 (ONS, 2013b). The Gross Value Added (GVA)[2] indicator is only 7% in total UK (100%) and ranked fourth from the bottom within English counties, with labour productivity only 89.6% compared to 100 as the UK baseline. The gross disposable household income in 2012 was £13819 compared to an average figure of £16251 for England, the second lowest amongst the English regions (ONS, 2013b). The gross weekly earnings for the full time adult employees was £465 compared to the UK median of £506 making the residents of the region one of the lowest paid group of employees in England. The manufacturing sector has a relatively high proportion of 15% in the total GVA compared to 10% for the total of the UK (see Figure 5.2).

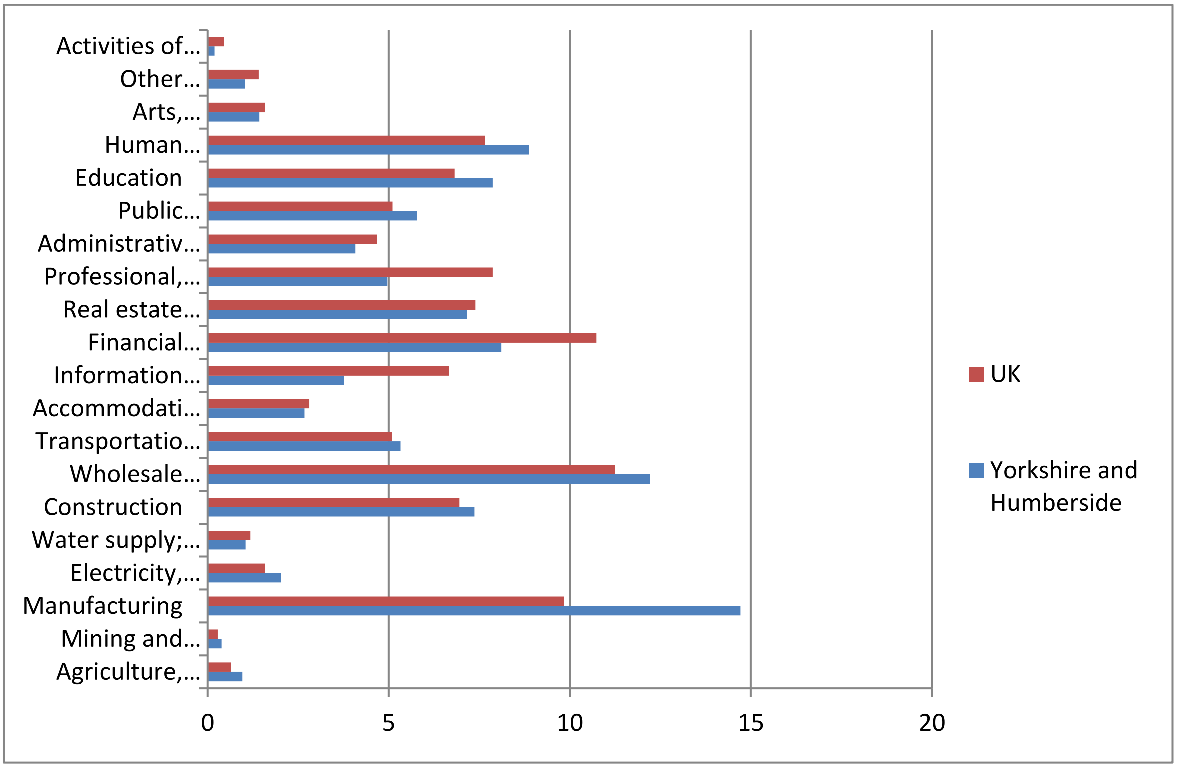


Figure 5.2. Yorkshire and Humberside Economy, 2013. Source: ONS, 2013a

Wholesale and retail trade, with health and social work activities, make a high contribution to the region's economy with 12% and 9% respectively, which is higher than the UK figures. The region is lacking in the provision of financial, professional, scientific and technical activities which are below the overall UK's figures. Distribution

[2]Gross Value Added (GVA) is the measure of the value of goods and services produced in the area, industry or sector of an economy. GVA is used to estimate GDP (Source: ONS, 2016)

of other industries is very similar to the rest of the UK. Consequently, the region has a higher than average proportion of adult population aged 16 to 64 employed in low level qualification occupations with 13% compared to 11% in the UK.

The Figure 5.3 demonstrates the distribution of population belonging to the highest social classes A and B in the region. The north of the region around the seaside resorts Whitby and the Yorkshire Moors, district of East Riding of Yorkshire and Yorkshire Dales have the highest concentration of professionals and middle class residents with up to 37% of this category living there. The areas around the seaside resort of Bridlington, large cities, e.g. Leeds, Sheffield and Kingston-Upon-Hull have 10% or less of AB classed residents living in these neighbourhoods.

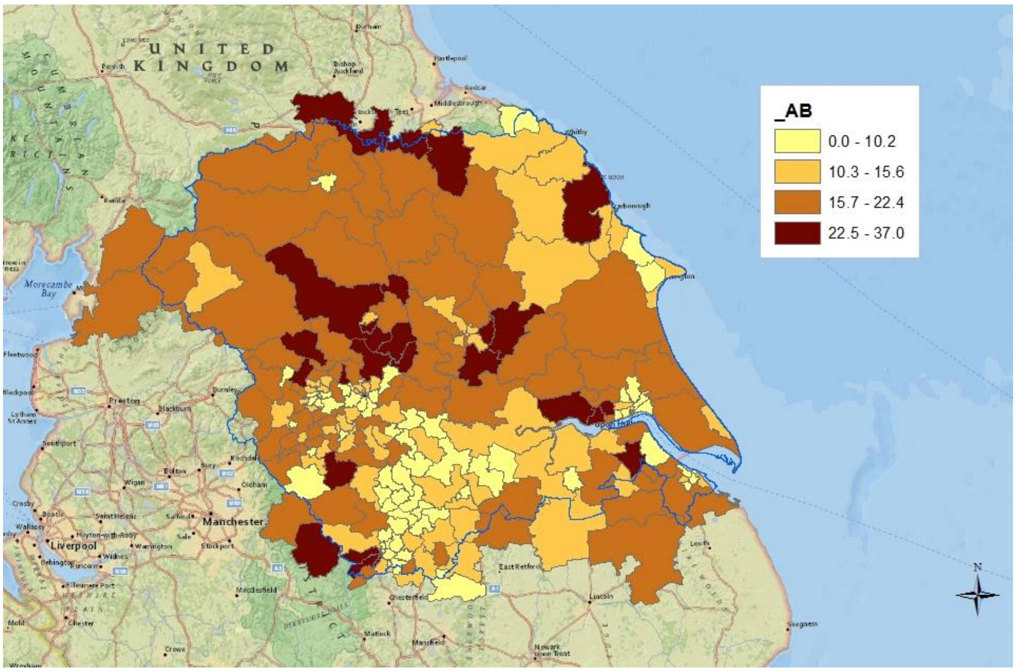


Figure 5.3. Distribution of AB social class population in the area. Source: ONS, 2013

In general, Yorkshire and Humberside is a very average region in the UK in comparison to geographic size, population density and birth rate. Moreover, the regions' population is mostly concentrated in urban areas as in other regions in the UK. However, it has distinctive differences in terms of the lowest productivity, many areas of high deprivation, a manufacture oriented economy and more land in national parks than any other English region.

5.3. Data Sources

To achieve the research aims and objectives this thesis uses secondary data. The data was provided from two reputable sources – CACI, the marketing agency and the major supermarket chain.

Section 5.3.1 provides an overview of CACI, a leading marketing agency and analyses the supplied data in terms of online grocery expenditure within various product categories Section 5.3.2 introduces the supermarket chain and their loyalty card scheme, Nectar, and provides an overview of the supplied data in terms of weekly online and face to face expenditure.

5.3.1. CACI data

CACI (Consolidated Analysis Centre, Incorporated) is an American based company specialising in information technologies and providing professional services to private and government organisations. Founded in 1962 the company opened the UK based CACI Limited in 1975 (CACI, 2016). Their services range from marketing, location planning, network services to technology solutions. CACI designs tailor made applications and tools in consumer segmentation, forecasting and retail catchments.

Using their own geodemographic profiling system (ACORN), CACI have produced estimates of e-commerce usage across different ACORN profile groups based on the Living Costs and Food Survey (LCFS) data, which covers about 50,000 respondents annually. CACI also have a set of control figures for online penetration, which are estimated based on ONS Retail Sales Index (RSI), industrial reports, and company annual reports. The Demographic data is based on 2012 population figures and the Acorn Population counts are based on 2013 figures.

The following data for the Yorkshire and Humberside region have been provided by CACI at the Postal Sector geography level for a total of 792 postal sectors.

1. Demographics within 47 parameters including Age, Gender, Ethnicity, Economic Activity, Social Grade and Family Structure. Table 5.2 provides summary of five major demographic groups and 47 demographic subgroups.

Table 5.2. Yorkshire and Humberside demographics

Age Bands	Ethnicity	Economic Activity	Social Grade	Family Structure
Age 0-15	Total	Persons 16-74	Persons 16-64	Families
Age 16-19	White	Econ active	AB	Couple family
Age 20-24	Mixed	Employee	C1	Lone parent family
Age 25-44	Asian	Self employed	C2	Male lone parent family
Age 45-64	Black	Unemployed	D	Female lone parent family
Age 65+	Other ethnicity	Fulltime student econ active	E	Family 0 dependent kid
Male age 0-15		Econ inactive		Family 1 dependent kid 0-4
Male age 16-19		Retired		Family 1 dependent kid 5-18
Male age 20-24		Other econ inactive		Family 2+ dependent kids youngest 0-4
Male age 25-44				Family 2+ dependent kids youngest 5-18
Male age 45-64				Lone parent 1 dependent kid 0-4
Male age 65+				Lone parent 1 dependent kid 5-18
				Lone parent 2+ dependent kids youngest 0-4
				Lone parent 2+ dependent kids youngest 5-18

2. Count of Population by Acorn category, type and group. This detailed classification includes 62 types, which are aggregated into 16 groups from A to P. The ACORN demographic classification has been described in Chapter 3.

3. List of ProVision stores (Grocery supermarkets), their store sizes (square foot) and X and Y co-ordinates for the total of 511 stores in Yorkshire and Humber.

4. Estimated weekly household expenditure by products for two retail channels - residential (face to face) and online. The classification COICOP (Classification of Individual Consumption by Purpose) is used to categorise the products. The grocery expenditure includes four COICOP groups – Food, Non-Alcoholic Beverages, Alcoholic Beverages and Cigarettes. The complete list of this classification is attached as Appendix B.

CACI Spend Estimates and Projections provide retail consumer spending patterns which are consistent with the Government National Statistics and calculated in collaboration with Cambridge Econometrics with the use of their Regionalised Multi sectoral Dynamic Model (MDM), a widely used model for the UK economy at the regional level (CACI, 2013). CACI's own local area models, together with MDM, produce forecasts for 92 product groups and four purchase types (comparison, convenience, motor fuel and services) for any geographical area in the UK. They use flat real-term projections for the retail forecasts with non-increasing level of spend in real terms. The four spend modes (residential, online, workers and tourists) reflect the total expenditure within each category with no double counting. The residential expenditure originates at customer home addresses. Online expenditure is defined as the purchase for which the payment is made online: 'Click and collect' transactions where the payment is made in a physical store are not classed as online expenditure. The online spend estimation is the total expenditure made online by people resident in each postal sector. These estimates are based on various open data sources including the Living Costs and Food Survey (LCF). The overall online penetration is based on ONS RSI figures and projections derived from the ONS RSI time series. CACI's researchers expect online penetration to continue to increase in the future at a linear rate until 2025 (CACI, 2013). CACI's UK spend estimation approach follows two phases. Firstly, product group prices and expenditure are calculated at the national and regional levels. Secondly, the data is disaggregated at the local level with model applications. CACI uses logistic and generalised linear models to predict buyer rates (the proportion of the population in a local area who will spend money on the product) and spend rates (the average expenditure among all consumers). Local area spend at the product group level is disaggregated by detailed product lines based on the Acorn classification. Figure 5.4 provides an overview of the CACI methodology.

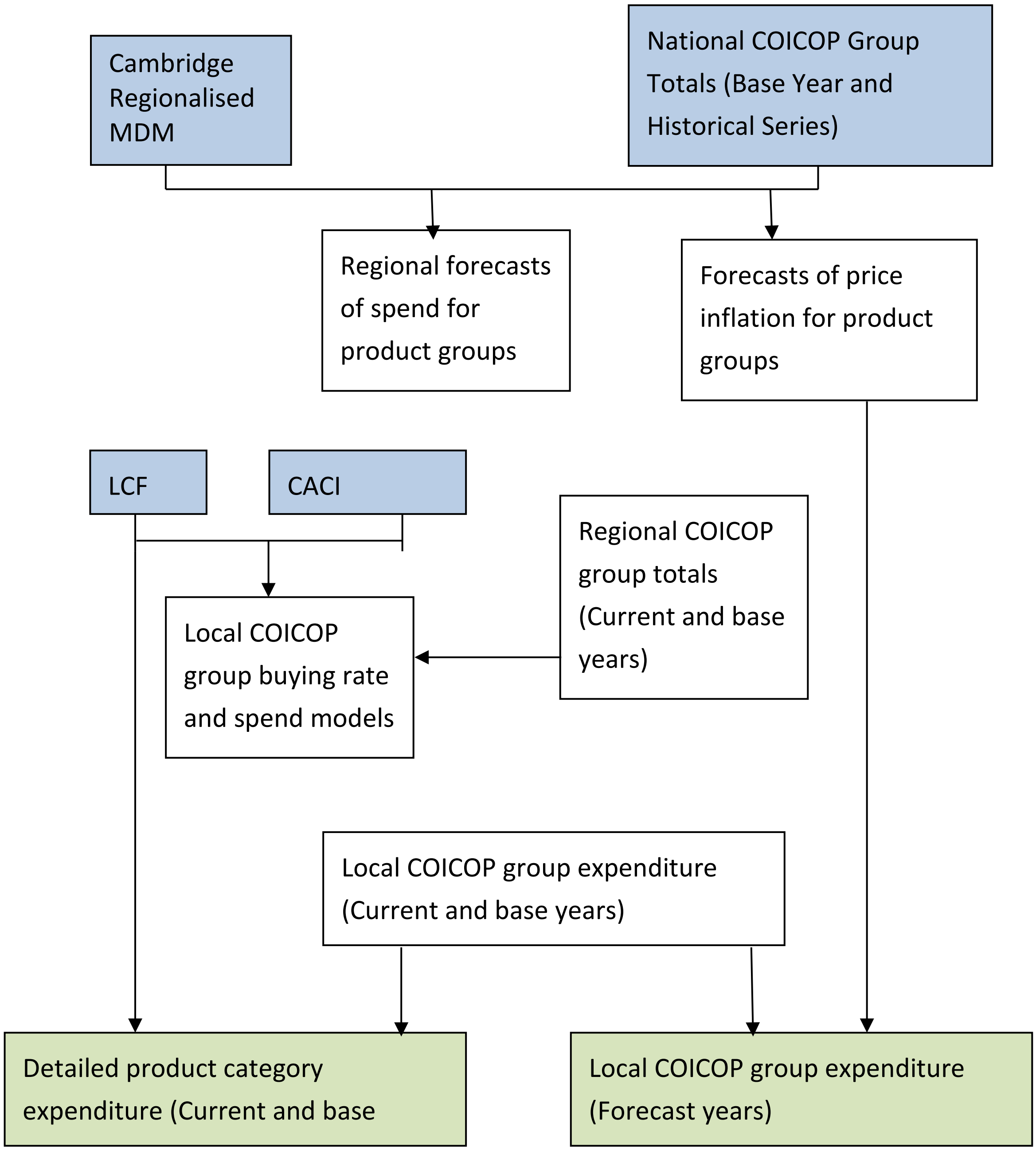


Figure 5.4. CACI methodology of Estimates and Projections. Source: CACI, 2013

The inputs (Cambridge MDM, COICOP Groups, LCF and CACI Area data) are indicated in blue and outputs are indicated in green.

Online estimates and projections are based on the ONS Retail Sales Index and the ONS index of Internet connectivity which are enhanced with information published by private companies (annual reports and trading statements from the major Internet retailers).

5.3.2. Sainsbury's position in the grocery sector and the supplied data

Sainsbury's is the second largest supermarket chain in the UK with 17% market share in the grocery market (Chapter 2). It was founded in 1869 by John James Sainsbury and his wife with the opening of their first shop in London. Their self-service groceries stores became very popular and the group grew very rapidly becoming the largest grocery retailer in 1922 and maintain its leadership until the 1990s when Tesco seized the leadership in the UK grocery market. In 2003 the American based ASDA propelled Sainsbury's back to third position in the grocery sector, although in 2014 the supermarket chain regained second place in the grocery market overtaking rival ASDA again (Ruddick, 2014). The historic success of Sainsbury's is due to its innovative approach and the core principle of a combination of high quality products at reasonable prices. Sainsbury's was the first grocery retailer to introduce own brand products with many matching the quality of national brands but at lower prices (J Sainsbury plc, 2015). The group has grown organically rather than through acquisitions and takeovers. The company's main focus has been on food while their competitors Tesco and ASDA have diversified into the clothing market and other non-food goods. Although, the company did invest into the do-it-yourself (DIY) market under the Homebase name with a supermarket format (The Independent, 2008). This venture was very successful with Sainsbury's finally selling the business with a twofold increase worth £969 million. In the 1980s Sainsbury's followed the trend in the grocery market by expanding into out of towns locations by opening large superstores. In the 1990s the company lost its dominance due to not expanding into new markets (clothing) and poor performances in the United States and not introducing a loyalty card scheme (unlike the closest competitor Tesco) until relatively recently. Moreover, its complex relationship with suppliers and inflexibility in price policy caused significant decrease in profits. In the late 1990s Sainsbury's reinvented itself by creating a new corporate identity with the new mission statement "Making life taste better" which later was replaced with the slogan "Try something new today" and finally in 2013 the current slogan "Live well for less" (J Sainsbury's, 2016). The company finally followed the trend and moved into new markets (clothing and banking) and developed new channels (convenience and online) and became very successful once again. Moreover, Sainsbury's expanded into new locations by taking over Somerfield and The Co-operative stores with extending their presence from the traditional south east locations of the UK into northern parts of the country, Scotland, Wales and Ireland (BBC, 2009). Moreover, the group relocated some of their major services and headquarters from London to the north of the country – Manchester, Lincoln and Coventry. Following a tradition of innovation, Sainsbury's have invested into renewable energy and converting food waste into bio-methane gas to generate electricity. As a result Sainsbury's store in the West Midlands became the first self-sufficient grocery store for energy in the UK (The Guardian, 2014). Despite its expansion into northern parts of the UK, the company dominates in the south east of the country with high concentrations of stores in the Swindon, Wolverhampton, Guildford, Redhill, Darford, South East London and Enfield postal areas (BBC, 2006). Currently, the company has 1312 stores across the UK with a total of almost 23 million sq ft floorspace (J Sainsbury plc, 2015). The comparative analysis with other major supermarkets has been outlined in Chapter 2.

This section analyses the data provided by the client on e-commerce sales across Yorkshire and Humberside in the UK. The data consists of the client online expenditure derived from the loyalty card scheme from 814000 unique customers for three months in 2013 which accounts for 15% of the total study area population of 5.3 million and could be considered as a substantial sample (ONS, 2015). The data for online channel usage could be considered as representative of all the client's online customers as almost 100% of client's customers use their loyalty card when buying online. The supermarket has a large presence in the area with over 130 stores ranging from small convenience stores (with sales area between 800 to 3000sq ft), supermarkets (with sales area of up to 25000sq ft) and hypermarkets (with over 60000sq ft sales area) (IGD, 2016). The data has been aggregated from individual post codes to the postal sector geography level of resolution.

Prior to analysis of the data it is necessary to provide a brief overview of the loyalty card scheme and its impact on the grocery retail industry. The loyalty card scheme is increasingly popular with half of UK households collecting points while shopping in stores and online, buying holidays, using financial services, etc. (J Sainsbury plc, 2016). The recent economic downturn boosted attractiveness of the loyalty schemes as shoppers became more 'savvy' and cautious in their spending patterns trying to maximise their money in terms of using cashback sites and various loyalty schemes.

The loyalty card scheme is a very effective marketing tool, increasing retailers' competitive advantage, helping to maintain market share and build relationship with customers (Burt et al, 2010). Moreover, the data collected with the assistance of the loyalty scheme is an invaluable source of information for the decision making process within the retail organisation, from stock management and staffing to the selection of site location. For example, Humby et al (2008) state that Tesco's introduction of the Clubcard scheme has changed the traditional grocery retailing operation with strategic and product development decisions now being made using extensive loyalty card scheme. These data provide detailed insights into consumer behaviour and their demographic profiles. Retailers can process information regarding their customers in terms of frequency of visits, amount spent, home or workplace locations and products preferences. Retailers realised the potential of the loyalty scheme data by creating consumer insight units which collect and analyse the data and providing outcomes across the business. The loyalty card scheme, from a straightforward marketing tool has become a very powerful strategic mechanism providing an insight into consumer trends, greater management of in store operations and helping to identify future vision of the business. For example, loyalty card schemes have helped grocery retailers to identify new areas of expansion, e.g. financial services.

Sainsbury's launched its Nectar card scheme in 2002. Customers are rewarded points for every £1 they spent in-store or online. To enhance its competitiveness Sainsbury's introduced the 'coupon at till' initiative where customers are rewarded with the coupons at the check outs for hundreds of branded and own products. The coupons are tailored to reach the target market of customers who will be interested in these offers. Sainsbury's claims that customers can earn points on half of the expenditure on household goods. Currently the Nectar loyalty card is the most popular loyalty card scheme with more participants (almost 17 million) than Tesco's Clubcard and Boots

Advantage Card. Its increasing popularity is due to the variety of its partners (14 in total) and 400 online retailers in the Nectar scheme. The Nectar scheme provides comprehensive data on customers' consumption and behaviour, although no retailer can achieve 100% loyalty scheme participation with 65 to 75% being the highest membership score (Humby et al, 2008). It was noted that some demographic groups (students) are not likely to participate in loyalty schemes and the amount of spend also effects usage of the loyalty card as customers tend not to use loyalty card for petty transactions (buying a sandwich or milk). Online channels have the highest uptake of loyalty card usage with almost 100% of customers registering their loyalty card during the transaction as customers expenditure is much higher compared to in store expenditure. Client estimates that 53% of their supermarket customers use Nectar cards compared to only 20% customers swiping their card while buying groceries at convenience stores.

The following data has been provided by the client for the study area with 814301 unique users at the output area geography level (17227 in total).

1. List and description of the groceries stores across the region with a total of 1753 stores. Below are the attributes of each store's data

Table 5.3. Stores' attributes

Store Reference	Salad Bar	Non Food SqFt
Open 24 Hr	Oriental Curry	Date of Last Investment
Cafe	Pizza Bar	Latest Investment Type
Home Shopping	Hot Food Counter	Z_Govt Region
Car Wash	Pharmacy	Petrol
Hand-Scanning	Lottery	Mezzanine
Bakery	Sales Area (SqFt)	Easting
Delicatessen	Sales Area Inc SqFt	Northing
Meat Counter	GM Sales SqFt	Location (Latitude)
Fish Counter	Clothing Sales SqFt	Location (Longitude)

In this research the grocery sales area (in square foot) has been used which were calculated using the total Sales Area (Sq Ft) and Non Food Sq Ft attributes. The Easting and Northing data allowed the locations of all grocery stores in the study area to be mapped and used for the calculations of the distance matrix (which will be described fully in Chapter 7).

2. Expenditure and frequency of transactions across three channels – Supermarkets, Convenience stores and online. Table 5.4 shows the sample of the data

Table 5.4. Customers' expenditure by Output area

Cust ID	OA	JS	JS Trans	Local Sales	Local Trans	Online Sales	Online Trans	Total Sales	Cust Score	Cust Type
1	E00058143	189.15	4					189.15	1	JS Only
2	E00067387	865.22	20					865.33	1	JS Only
3	E00054064	244.36	7					255.36	1	JS Only
4	E00053906	508.33	25	17.34	2			525.67	4	JS& JS Local

A summary of the data with expenditure across the three channels derived from the loyalty card scheme is presented in the Table 5.5

Table 5.5. Client's channel use in three month period

Channel	Sales, £	Sales %	Customers	Customers %	Transactions	Transactions %
JS Only	136,487,123	64.4	525628	64.5	4,472,492	57.9
JS & JS Local	46,905,145	22.1	155182	19.1	2,172,528	28.1
JS & Online	8,597,755	4.1	11796	1.4	145,117	1.9
Online only	8,387,115	4.0	28205	3.5	75,110	1.0
JS Local Only	7,135,071	3.4	87141	10.7	755,608	9.8
All 3 Channels	3,800,963	1.8	5160	0.6	92,208	1.2
JS Local & Online	516,802	0.2	1189	0.1	13,379	0.2
Total	**211,829,975**	**100**	**814301**	**100**	**7,726,442**	**100.0**

The data shows that the most popular channel for groceries is in-store with over half of all grocery transactions completed at the supermarkets, with over 60% of customers using this channel only. Almost 20% of all customers use both physical channels – supermarkets (JS) and small convenience stores (JS Local). The third most popular channel is convenience stores with 10% of client's customers buying their groceries at local convenience stores. The online channel is not as popular, as only 3.5% of all client's shoppers prefer to buy their groceries via the website only with 1% of all grocery transactions completed online. In spite of the increasing popularity of multichannel shopping the real data shows that less than one percent of customers use all three channels to buy their groceries. The least unlikely combination of channels is convenience and online with 0.2% transactions via these channels. Overall, the most popular channel is supermarkets (JS) with almost 86% of all customers using this channel. Table 5.6 shows that almost a third of all customers buy their grocery at the convenience stores and almost 6% of all client' customers use online channel to buy their groceries

Table 5.6. The grocery expenditure by channel

Channel	Customers	Customers, %
Local	248672	30.5
JS	697766	85.7
Online	46350	5.7
Total	*814301*	

It is useful to explore the patterns of online expenditure in comparison to two major physical grocery retailing channels – convenience stores and supermarkets. The Table 5.7 shows the average expenditure per transaction across all three channels. Customers are likely to spend higher amounts online with almost £114 per transaction compared to £8.85 per transaction at convenience stores. The average expenditure per visit at the supermarkets is £30.85. The data corresponds with the national average weekly grocery expenditure of £80 per household (DEFRA, 2015).

Table 5.7. The grocery expenditure and transactions by channel

Channel	Sales, £	Transactions	Average Transaction, £
JS	182,308,914.16	5909779	30.85
JS Local	14,951,033.77	1688698	8.85
Online	14,570,026.90	127965	113.86

To reflect only household expenditure and thus eliminate likely business transactions, online expenditure of over £200 per transaction has been excluded from the analysis. Table 5.8 shows that for the period of three months in 2012 almost half of all online transactions were made just once, 5% of the customers purchased grocery online monthly and only 2% of the customers were frequent online shoppers. The percentage of the customers who shop more regularly than once a week significantly decreases from 2% to 0.3% of all transactions with the frequency of 13 in 12 weeks. The high average expenditure of £104 per online transaction is perhaps due to the supermarket's offer of free delivery on orders of over £100.

Table 5.8. Frequency of online transactions for the period of 12 weeks

Frequency	%	Average expenditure per transaction, £
1	53.4	90.3
2	15.0	101.8
3	8.2	106.5
4	5.1	108.9
5	3.6	110.3
6	3.0	109.8
7	2.0	111.0
8	1.7	108.0
9	1.6	109.0
10	1.7	109.6
11	2.0	111.0
12	2.1	109.3
13	0.3	108.5
14	0.1	103.2
15	0.1	96.7
16	0.0	116.8
17	0.0	106.3
18	0.0	96.7
19	0.0	198.9
20	0.0	82.4
21	0.0	91.4
22	0.0	114.3
23	0.0	72.7
24	0.0	82.7
25	0.0	87.5
27	0.0	66.7
32	0.0	107.9
	100	**104.4**

Analysing other channels of retail distribution – convenience and supermarket sales - there is a substantial variation in comparison to online sales. Consumers spend substantially less in supermarkets and local stores with £18.76 and £8.71 per transaction respectively, although, these two channels are used more frequently with up to 199 transactions in 12 weeks in supermarkets and 236 times in local convenience stores (compared to 32 transactions made via online channel).

3. Online expenditure by product categories including food and non-food (52 in total) within Output Area geography for three month period which were compressed into 12 major groceries categories. The total results are presented in Table 5.9 below.

Table 5.9. Client's twelve weeks online expenditure by product categories

Product Categories	Online Sales, £	%
ALCOHOL	1519371.08	12
TOBACCO	453268.19	4
MEAT/FISH/SEAFOOD	1227691.27	10
FRUIT/VEG	3257001.3	26
DAIRY/CHEESE	1384713.95	11
DESERTS/BREAD	854302.21	7
SOFT DRINKS	585249.46	5
STORE CUPBOARD BASICS/CANS/SUGAR/RICE/HOT BEVERAGES	756490.64	6
CONFECTIONARY/SNACKS	754721.41	6
ETHNIC	13193.31	0
CONVENIENCE/DELI	1061548.82	8
FROZEN	783496.94	6
Total	**12651048.58**	**100**

The total for twelve major grocery categories represents over 80% of the total supermarket online expenditure which indicates that the majority of the expenditure is within food categories. Data in the second paragraph (expenditure across three channels) includes all 52 products categories. Consequently, the data for the total online channel is higher compared to the data in Table 5.8.

4. The actual sales data for 131 client's stores in the study area which is derived from the Point of Sales (POS) represents accurate expenditure levels at the supermarket stores. This data was used to help calibrate the estimated data. The overall supermarket weekly expenditure in 2012 was £24.4 million compared to £17.7 million (Table 5.5) or 28% underestimation through the loyalty card scheme alone. These differences in data were considered in model estimation and calibration (see Chapter 8).

Considering all the above facts the received data is considered to be substantial and representative corresponding well with national statistical data. Client's extensive sample data (with over 800 000 unique users) represents 15% of the total population in Yorkshire and Humberside (ONS, 2016). The data is equally distributed between rural and urban locations. Although, some bias is expected as customers may not provide the accurate data. CACI in collaboration with Cambridge econometrics, uses an advanced methodology to produce expenditure forecasts. Both sets of data have a high correlation in terms of online expenditure and consequently the obtained data can be considered as valid and representative. Although own demand estimates will be produced in Chapter 6 to check the validity of the data. These two companies do not exist in academia and this research has a rare opportunity to use both sets of data (estimated and actual) to calibrate the model. Although CACI has a comprehensive estimation methodology the online expenditure data needs to be tested and replicated. In Chapter 8 the model attempts to replicate the CACI online estimates using similar technique with further model calibration against the actual online expenditure data provided by client. The more detailed analysis of the actual and estimated data from CACI and supermarket chain with regards to demographics and demand estimation in the study area will be explored in the next chapter.

Chapter 6: Building demand estimates for e-commerce purchases

6.1. Introduction

The demand for on-line buying has been estimated in a small number of previous empirical studies, i.e. (Thompson et al, 2013). The key geodemographic characteristics, i.e. age, ethnicity, social class and gender will produce different estimates of need. The demographic characteristics of the population drive these spatial variations in demand, as different segments of the population require different services. This demand can be spatially and statistically analysed through examining on-line buying behaviour in terms of purchased products and customers' demographic characteristics. Few studies to date have evaluated variations in on-line activity relating to demand by small area populations. Geodemographics analysis will allow to establish a profile of on-line users. This chapter will build demand layers for e-commerce expenditure based on key factors deemed important in the literature, combined with survey data. Finally the actual data of on-line expenditure provided by the client will be examined. First, Section 6.2 attempts to build on-line demand based on major demographic characteristics of on-line grocery customers as identified in Chapter 3. Secondly, Section 6.3 provides on-line demand estimation for the study area with comparison to CACI data. Furthermore, the actual on-line expenditure will be analysed in terms of customers' locations and expenditure patterns (section 6.4). Finally, Section 6.5 analyses on-line expenditure within product categories using both estimated and actual data from the following sources – Supermarket, DEFRA, Mintel and CACI.

6.2. Building demand based on literature review

On-line demand is based on the total residential demand which will be explored in more detail in Chapter 7 as it forms an essential part of any spatial interaction model. The demand side modelling aims to make use of regional, county or finer scale spatial data (Office of National Statistics and CACI) so that it can be replicated for any region or area within the UK.

Chapter 2 outlined the two principal theories (efficiency theory and diffusion of innovation) which indicate that accessibility and demographic characteristics of on-line customers are the two major factors of building a demand layer for on-line expenditure estimation. Chapter 3 highlighted that age, social class and gender are the major demographic characteristics of online customers. Based on the data outlined in Figure 3.14 and Figure 3.15, showing the distribution of customers who buy their groceries regularly on-line, the following estimation maps have been produced using the demographics data obtained from CACI by postal sector geography (as described in Section 3.2.1.). The demand has been calculated as the percentage of total population per postal sector. First, Figure 6.1 shows the distribution of on-line customers based on age characteristics.

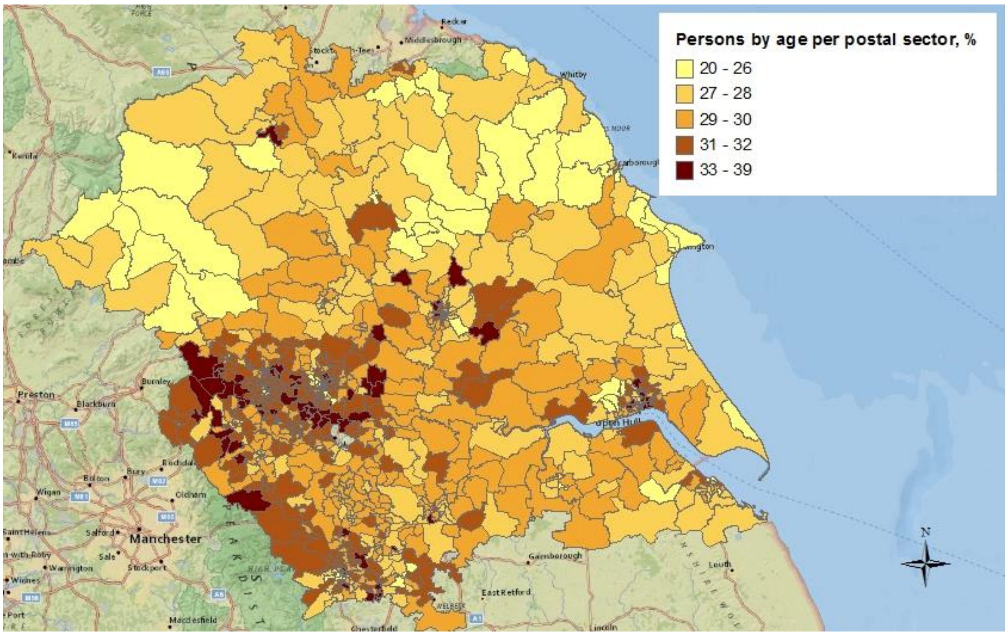


Figure 6.1. Estimation of on-line customers using age

The highest concentration of on-line customers (of up to 39%) is around large cities – Leeds, Kingston-upon-Hull, York and Sheffield. The estimation of demand for on-line groceries based on age confirms the diffusion of innovation theory with the assumption that the number of on-line customers is likely to be high for young professionals living in the cities. Areas with a more elderly population, including many rural areas in the north of the study area, have the lowest estimated distribution of on-line customers (as low as 20%). The next maps explore the estimated demand based on other customer's characteristics, showing how different demographic analysis can produce alternative estimated distributions of on-line grocery customers.

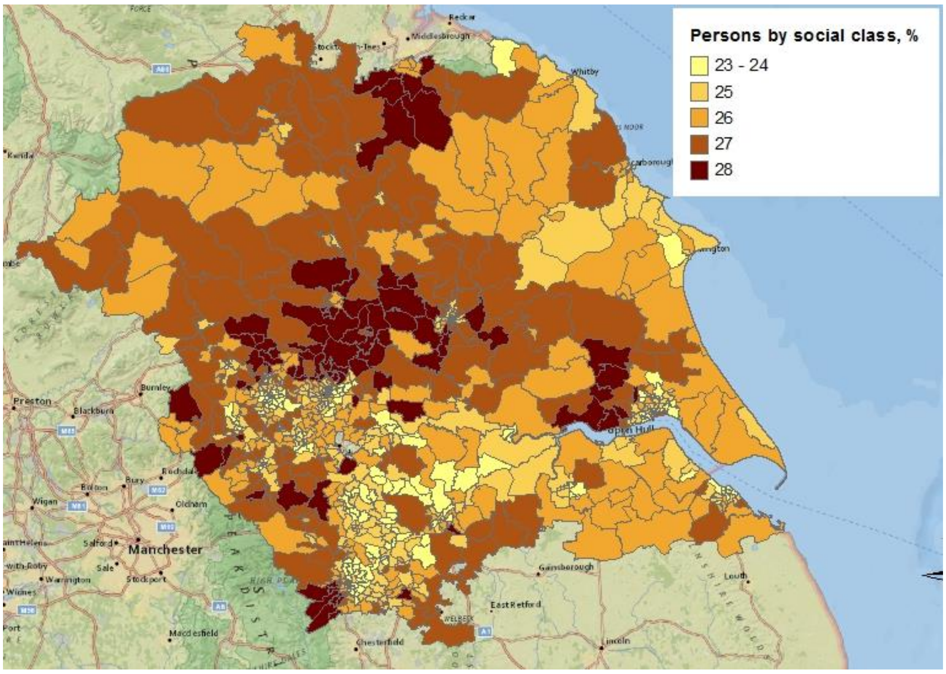


Figure 6.2. Estimated distribution of on-line customers using social class data

Figure 6.2 shows the distribution of on-line customers based on social class data. The highest concentration of on-line customers(of up to 28%) can be seen in the affluent areas of York, Leeds and Scarborough and in the northern, more rural parts of the study area. The lowest expected demand is in the less affluent areas, especially around cities such as Leeds and Bradford and in the southern parts of the study area.

The efficiency theory states that people living in more rural areas with -more restricted accessibility to grocery stores are more likely to use online channels to buy their groceries. Figure 6.3 shows the estimated distribution of on-line customers based on the population density of individual postal sectors. The data was calculated using the data from Eurostat (2013) outlined in Table 3.3 with the estimated distribution of on-line grocery customers calculated as follows:

1. Individuals living in densely-populated areas (at least 500 inhabitants/km²) – 18%
2. Individuals living in intermediate urbanised areas (between 100 and 499 inhabitants/km²) – 20%

3. Individuals living in sparsely populated areas (less than 100 inhabitants/km²) – 23%

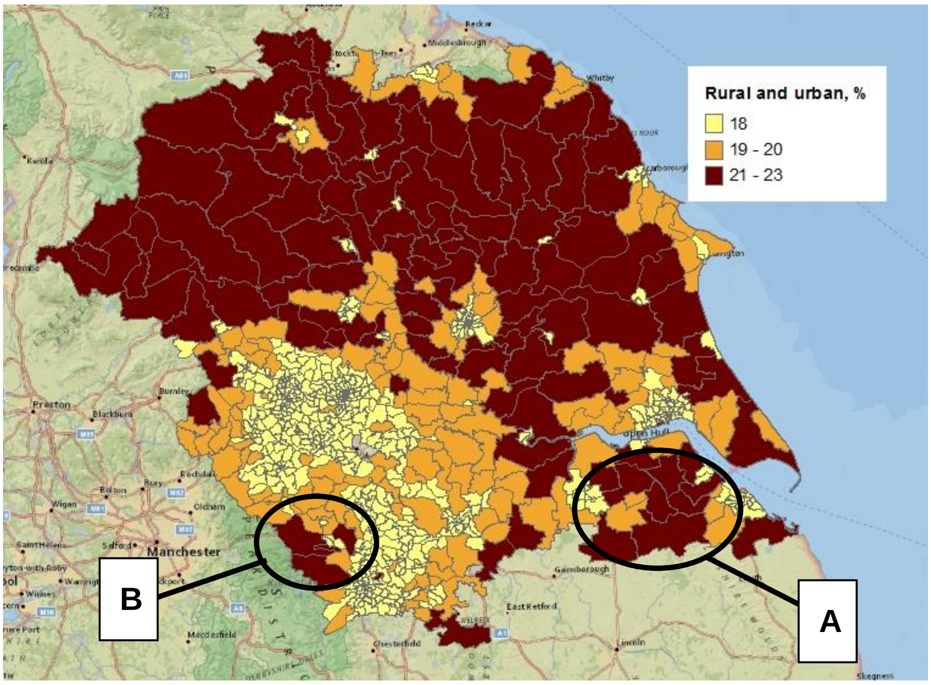


Figure 6.3. Distribution of online customers based on population density

Due the nature of the study area the largest proportion of on-line customers are located in northern rural areas (see Chapter 5 also). The highest concentration of on-line customers(of up to 23%) is in the north of the study area but also in rural parts of North Lincolnshire and Sheffield (Areas A and B).

Family composition is another important factor in determining the profile of on-line grocery shoppers. Research by the Institute of Grocery Distribution established that 32% of families with children under 5 buy their groceries on-line compared to 17% of households without children (IGD.com, 2012).

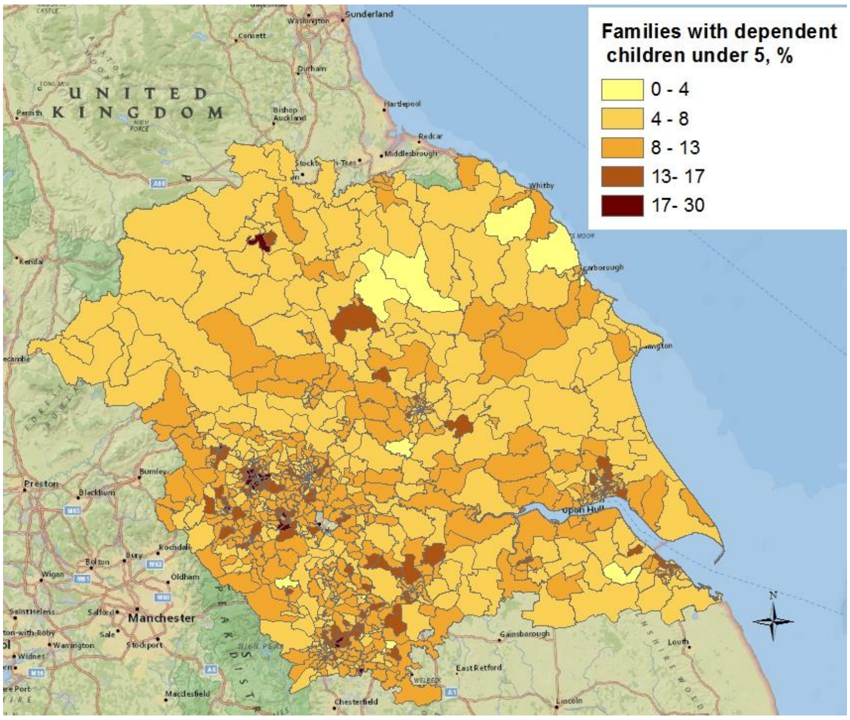


Figure 6.4. Distribution of families with dependent children under 5 years old

Figure 6.4 shows possible locations of on-line customers based on this family composition characteristic. The highest demand for on-line grocery would be expected to be in the areas of North Humberside, in certain urban areas of Leeds, Sheffield and Harrogate.

Comparing these four maps it could be surmised that the highest demand for on-line groceries is expected in the following areas:

1. Urban areas with high concentrations of young people and families with children under 5 years of age
2. Rural and more affluent areas with high concentrations of population within the A and B social classes

The next section will explore the demand for on-line grocery channel using CACI estimated data in comparison to own estimates of online expenditure in the study area.

6.3. Online Demand Estimation

As was described in Chapter 5, CACI provided estimated data for residential and on-line demand for groceries at the postal sector geography level of resolution. Using their own geodemographic profiling system (ACORN), CACI have produced estimates of e-commerce usage across different ACORN profile groups based on the Living Costs and Food Survey (LCFS) data, which covers about 50,000 respondents annually. Based on that data, Figure 6.5 shows the estimate of e-commerce usage for groceries across the Yorkshire and Humberside UK region by total on-line sales (Figure 6.5a) and by market share of total estimated grocery expenditure (Figure 6.5b).

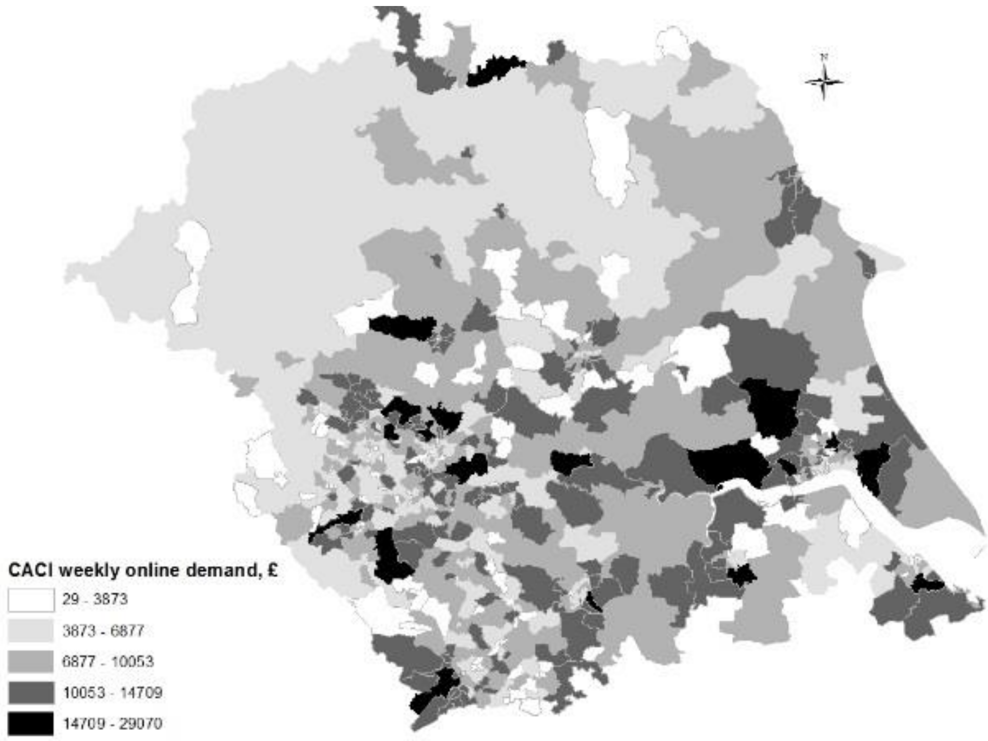


Figure 6.5a. Estimated weekly demand of weekly online groceries

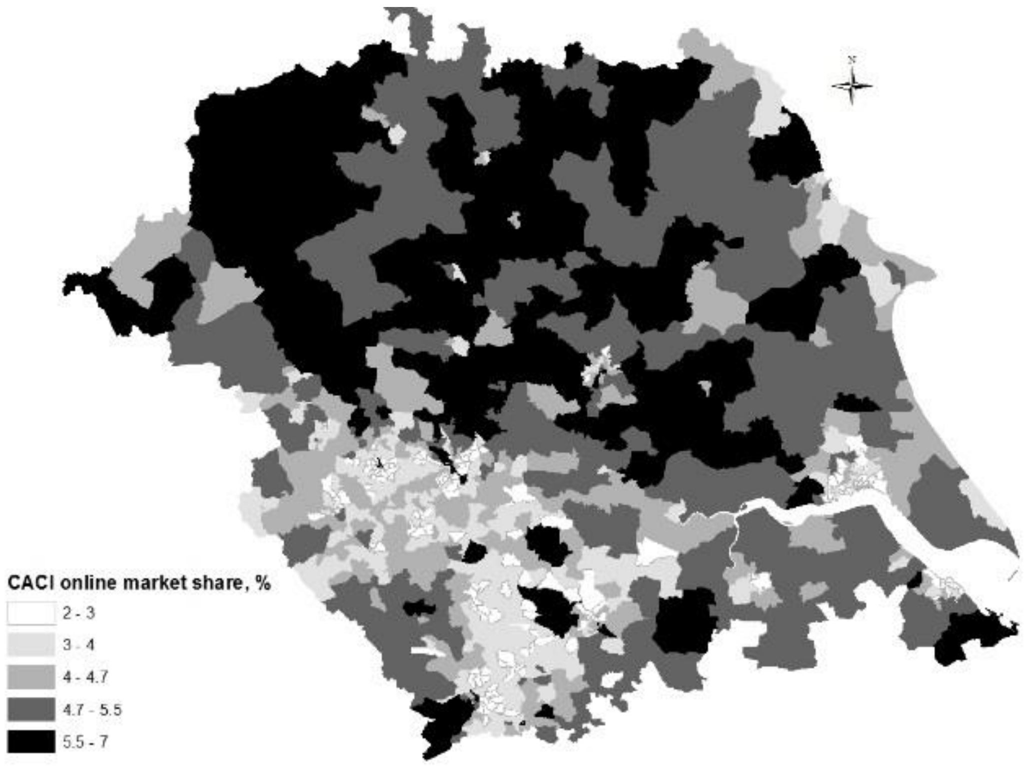


Figure 6.5b. Online market share in total grocery expenditure

According to the CACI data the largest demand for on-line groceries occurs in the more urbanised and suburban areas with the weekly on-line expenditure of up to £29k per postal sector around metropolitan boroughs, e.g. Leeds and Sheffield in the south west of the region and unitary authorities and large cities, e.g. Kingston-upon-Hull and Grimsby in the south east of the region and York and Scarborough in the northern part of the study area (Figure 6.5a). The distribution of the on-line demand corresponds with the distribution of population density within the region with the least populated areas in the northern part of the region having a weekly on-line expenditure as low as £29 per postal sector. Due to this fact the on-line market share of total grocery expenditure is more useful indicator in analysing the spatial distribution of on-line expenditure across the study area. Figure 6.5b demonstrates the different patterns of estimated on-line expenditure across the region in comparison to Figure 6.5a, with the rural and less populated northern areas having the largest market share of up to 7% and the urban districts largely generating smaller market shares of on-line expenditure. Interestingly, CACI estimations of the on-line share are higher compared to the national data of 4.4% (IGD, 2016), although, on-line share estimates are the same as the actual on-line share of 7% (see Table 6.4).

To use this data in subsequent analysis (see later chapters) it is important to verify the CACI data by estimating own on-line expenditure data. This is done using the LCFS data for 2012 of weekly household on-line grocery household expenditure disaggregated by social class.

To estimate household expenditure at a postal sector level, the following data will be applied:

- Number of households at postal sector level living in the region.
- The number of households was multiplied by average weekly expenditure taken from the Regional Family Spending Survey in England.

The average household expenditure is based on the classification COICOP (Classification of Individual Consumption by Purpose) which consists of 12 major categories. Table 6.1 represents the average household expenditure for 2013 of Yorkshire and Humberside region.

Table 6.1. Household expenditure in Yorkshire and Humberside, 2012-2013

No.	Expenditure article	£
1	Food and non-alcoholic drinks	54.10
2	Alcoholic drinks and tobacco	12.90
3	Clothing and footwear	18.90
4	Housing (net), fuel and power	53.20
5	Household goods and services	27.40
6	Health	4.40
7	Transport	54
8	Communication	11.10
9	Recreation and culture	57.20
10	Education	4.50
11	Restaurants and hotels	35.10
12	Miscellaneous goods and services	31.30
	Total	422.70

Source: ONS, Family Spending 2013.

The average weekly expenditure in the region of £422.70 in 2013was one of the lowest in the UK compared with the London area of £571.60 and the average UK household of £482.10 (ONS, 2013). The highest expenditure was on recreation and culture at £57.20 a week. In terms of food and non-alcoholic drinks expenditure, households in the study area spend less compared to the UK average of £56.80 a week, although, expenditure on alcoholic drinks and tobacco is slightly more when compared to the national average weekly expenditure of £12.60. To build the demand estimates for residential grocery expenditure the combined expenditure on food, non-alcoholic drinks, alcoholic drinks and tobacco per household (£67 in total) was multiplied by the number of households per postal sector in the study area (2.2million in total). The total grocery demand for the study area is £148million which corresponds with CACI data of £149 million. To calculate demand estimation for the on-line channel the national average figure of 4.4% was applied. Furthermore, based on the conclusions from the previous chapters (Chapter 2 and 3) stating that customers within higher social class are more likely to use on-line channels to buy their groceries, on-line grocery demand has been disaggregated using NRS classification by four social grades (AB, C1, C2 and D) across the 791 postal sectors. CACI data of the distribution of households by social class in the study area was applied which was compared to the national statistics data from Neighbourhood Statistics Census Data for 2013. The data was disaggregated based on the assumption that areas with higher concentration of customers within social classes AB and C1 will have a higher above national average on-line share of 6% and 5% correspondingly. In contrast, areas with higher number of population within lower social grades C1, D and E will generate lesson-line expenditure, with shares of 3% and 2% respectively. In postal sectors where the majority of households belong to C1 social grade, online grocery share equals 4%. Table 6.2 presents the estimated total expenditure on on-line groceries in the study area in comparison to the CACI data.

Table 6.2. Weekly Online demand for Yorkshire and Humberside

Social Grade	Households	%	Online Demand, £	%	CACI Demand, £
AB	499239	23	2006940.8	31.0	1871254.4
C1	838203	38	2807980.1	43.4	2326363.8
C2	712666	32	1432458.7	22.1	1640688.1
D	168255	8	225461.7	3.5	294647.1
Total	**2218363**	**100**	**6472841.2**	**100.0**	**6132953.4**

When mapped (Figure 6.6) the expenditure estimates show very similar patterns of on-line demand estimation provided by CACI (Figure 6.5a)

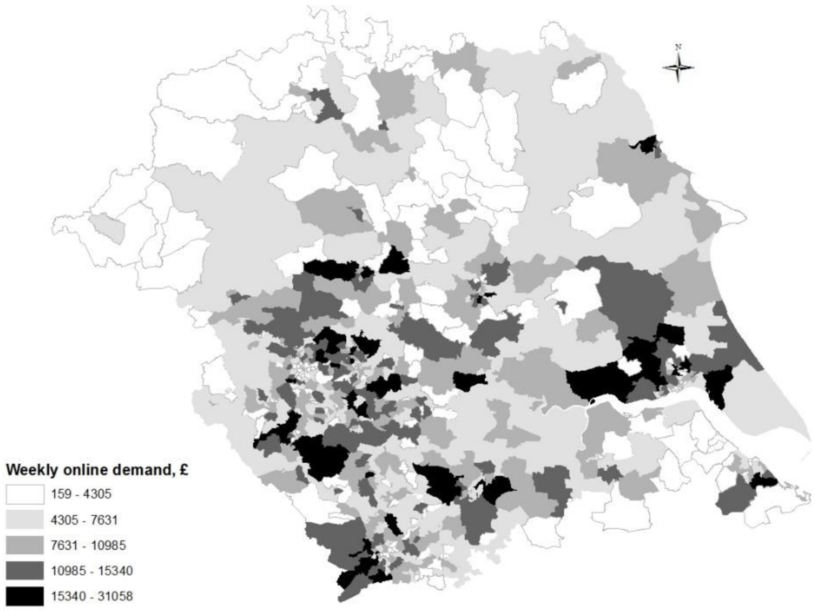


Figure 6.6. Estimated demand for e-commerce for Yorkshire and Humberside region by the author

The correlation between online estimated demand and CACI data is 82%.

The CACI and above described technique in estimating on-line demand are very similar and based on the socio demographic characteristics of the typical on-line customer. The CACI data is more up to date and has been used in other studies of e-commerce by the company itself. Given the fact that own estimates show similar patterns the CACI data can be applied in later modelling chapters with greater confidence. The next section will explore the demand for on-line grocery channel based on actual expenditure data.

6.4. Supermarket Chain Data

Chapter 4 reviewed some major studies exploring the geodemographics of e-commerce usage from various UK consumer surveys. This section analyses data provided by a major UK grocery retailer on e-commerce sales across Yorkshire and Humberside. As noted in Chapter 5, the loyalty card scheme has become a very important tool within the grocery market, generating data on customers' behaviour and allowing grocery retailers to make more successful strategic decisions. Chapter 5 identified various information which can be derived from the loyalty card scheme, e.g. customers' locations and their demographics characteristics. The Nectar card provides a wide spectre of information about grocery expenditure which is linked to unique anonymous customers (to protect their identities). The data allows to identify customer spatial origin, spending patterns and geodemographics characteristics. A total of 12 weeks data in 2012 were obtained for three different channels. Each record contains a unique customer ID number (linked to the customer loyalty card), transaction frequency and value, and customers' locations at output area geography based on customer's registered address when they signed for a loyalty card scheme (self-proclaimed home address).

6.4.1. Demographic characteristics of online customer within gender and age

The loyalty card data contains information about client customers' age and gender derived from the data provided by the customers at the point of registration with their date of birth and gender. The data which seemed unrealistic (i.e showing a customers' age under 16 and over 95 were disregarded). Total records of 781141 out of 814301, or 96%, were taken into account in calculating the distribution of customers. The date of birth data was aggregated into six age categories to comply with the CACI demographic data and national statistics records. The distribution between the genders was equal, with 51% males and 49% females. The distribution of customers within age and gender is presented in Table 6.3

Table 6.3. Distribution of client customers by gender and age

Female/Age	JS Transactions	Local Transactions	Online Transactions
18-24	4.5	5.3	4.4
25-44	30.6	33.4	30.7
45-59	31.4	30.4	31.5
60-64	9.6	9	9.3
65-75	16.2	15	16.3
over_75	7.8	7	7.7
	100	100	100
Male/Age	JS Transactions	Local Transactions	Online Transactions
18-24	5.1	6.4	5.8
25-44	31.9	34.8	34.4
45-59	30.9	29.8	30.3
60-64	9.3	8.7	9
65-75	15.6	14.2	14.6
over_75	7.2	6.1	6
	100	100	100

Table 6.3 shows that the most enthusiastic client's on-line customers are young males aged 25 to 44 with over 34% of total on-line transactions, although, this demographic category also has the highest rate for using the other two channels – Supermarkets (JS) and Convenience (local) stores with over 30%. Female on-line customers are most likely to be aged 45-59 and are the second most active category in on-line grocery shopping. However, this group also has the highest expenditure in the other two physical channels with 31% and 30% correspondingly. The least enthusiastic on-line customers are young people aged between18-24, especially females with the lowest on-line share of 4%. The older population aged over 65 are also unlikely to buy their groceries on-line, especially males who have the lower on-line rate of 6% compared to females of almost 8%.

Overall there is a conclusion that the most active on-line shoppers are males aged 25 to 44 and females aged 45 to 59, although there is no significant variation of expenditure across all three channels within these demographic categories. Moreover, the distribution of client's customers reflects the national distribution of population and study area in particular which was described in Chapter 5 (see Table 5.1). Statistically there was no evidence of a strong relationship between demographic characteristics (age and gender) and on-line expenditure. The next section will analyse client's on-line customers using multivariable demographic analysis which offers a more comprehensive demographic analysis combining various demographic characteristics which were described in Chapter 3.

6.4.2. Distribution of actual on-line sales

As described in Chapter 5 client's supplied data at the output area geography level. The data was aggregated to the postal sector geography level in order to compare to CACI data which uses postal sector geography.

Figure 6.7 shows the distribution of e-commerce sales across the region for supermarket chain by total sales and Figure 6.8 demonstrates share of on-line expenditure compared to total client's revenue. The average market share value for e-commerce is 10.3% of all sales which is more than double compared to the national figure of 4.4% of the total grocery market (IGD.com, 2016). The attractiveness of the on-line channel among client's customers is partly explained by the fact that they are more likely to belong to the higher social groups ABC1, who are also more likely to use the on-line channel.

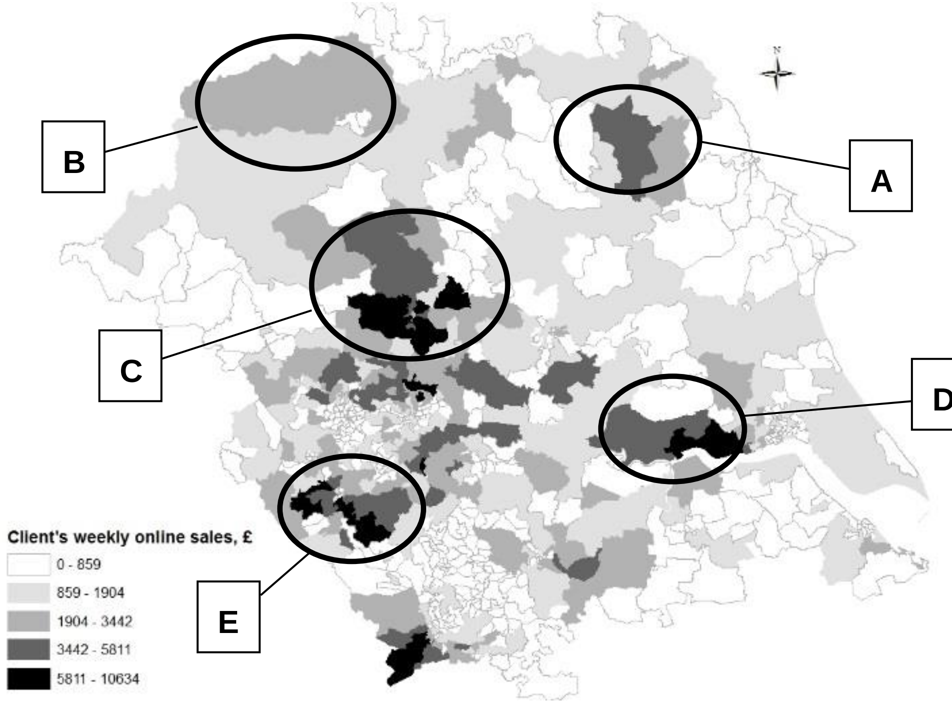


Figure 6.7. Client's weekly on-line revenue by postal sector

In terms of weekly on-line expenditure the highest on-line sales can be seen to be generated in rural areas in the northern parts of study area (Areas A and B), affluent suburban areas of Harrogate, Leeds, Kingston-upon-Hull (Areas C and D) and Sheffield (Area E). Figure 6.8 shows the distribution of on-line sales as a percentage of all estimated expenditure.

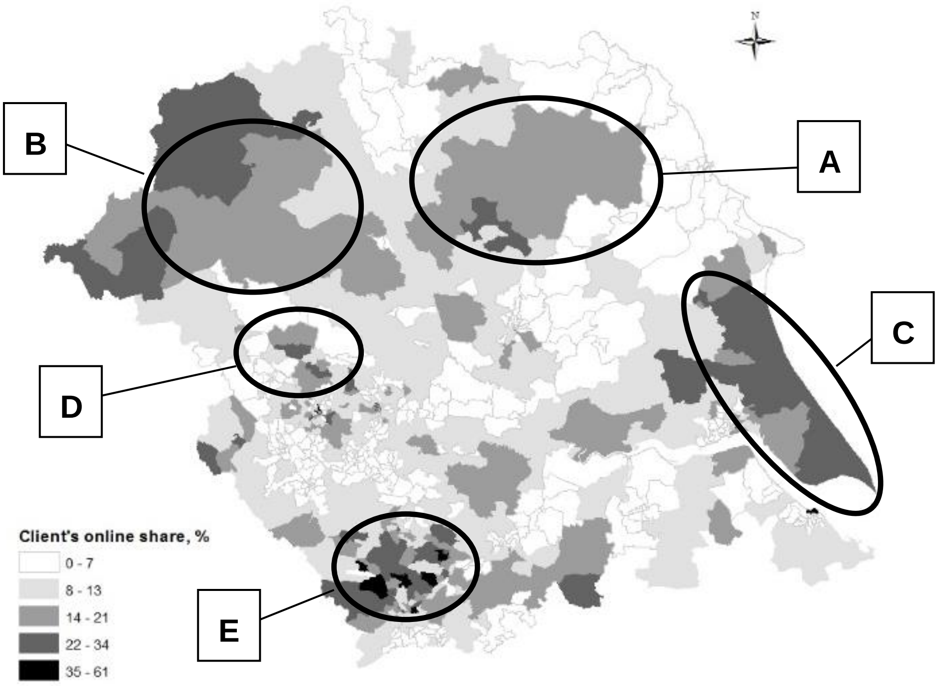


Figure 6.8. Client's weekly on-line market share

Figure 6.8 shows similar patterns compared to Figure 6.7, although the following tendencies are more evident:

1. On-line expenditure is more likely to be generated in rural areas with an on-line share as high as 61% (areas A, B and C)
2. On-line expenditure is higher in more affluent urban areas (i.e. areas D and E).

The more detailed analysis of client's on-line expenditure in terms also of accessibility will be explored below and further in Chapter 8.

Does these spatial patterns make sense which are observed in these maps? It would be expected areas with higher affluence and wealth to be important drivers of e-commerce activity. To explore this hypothesis the client's data was analysed by geodemographic classification, using the ACORN categories again here. The hypothesis seems to be borne out with the highest average on-line share of 13% within the affluent Rising Prosperity category (Acorn 2) and the lowest of 8% within the considerably less prosperous Urban Adversity category (Acorn 5) within ACORN demographic classification (Table 6.4). Moreover, client is the favourite supermarket of customers within Rising Prosperity category with an index of 143 compared to average national data (Table 3.3). Interestingly, customers within this Acorn category continue to be enthusiastic on-line shoppers despite the close proximity of a client supermarket with an average distance of 1.3km from the postal sector (centroids) to the nearest client's supermarket with grocery floorspace of over 3500 sq ft. The concept of distance and its relationship between on-line will be explored in more detail in Chapter 8.

Table 6.4. Online expenditure among ACORN demographic groups

ACORN	Postal Sectors	Online, £	Total, £	Online Share, %	Total Share, %	Average, %
1	178	490041	6219877	8	35	10.2
2	19	11188	156418	7	1	13
3	262	417108	6273948	7	36	10.2
4	212	205985	3304013	6	19	11.1
5	120	89745	1698242	5	10	8.8
Total	791	1214067	17652498	7	100	10.3

Overall, across all Acorn categories there is a tendency that the on-line share gradually decreases from 8% within the wealthy Affluent Achievers category to 5% within the least affluent Urban Adversity Acorn group.

Moreover, the analysis of on-line expenditure within OAC demographic groups shows similar results with more affluent demographic categories – Countryside, Prospering Suburbs and Typical Traits having higher percentages of total on-line expenditure with 18%, 32% and 25% correspondingly compared to 5% and 2% of total on-line expenditure within Constrained by Circumstances and Multicultural Communities.

The multivariable analysis of ONS data in comparison to actual sales data from the client is presented in Figure 6.9

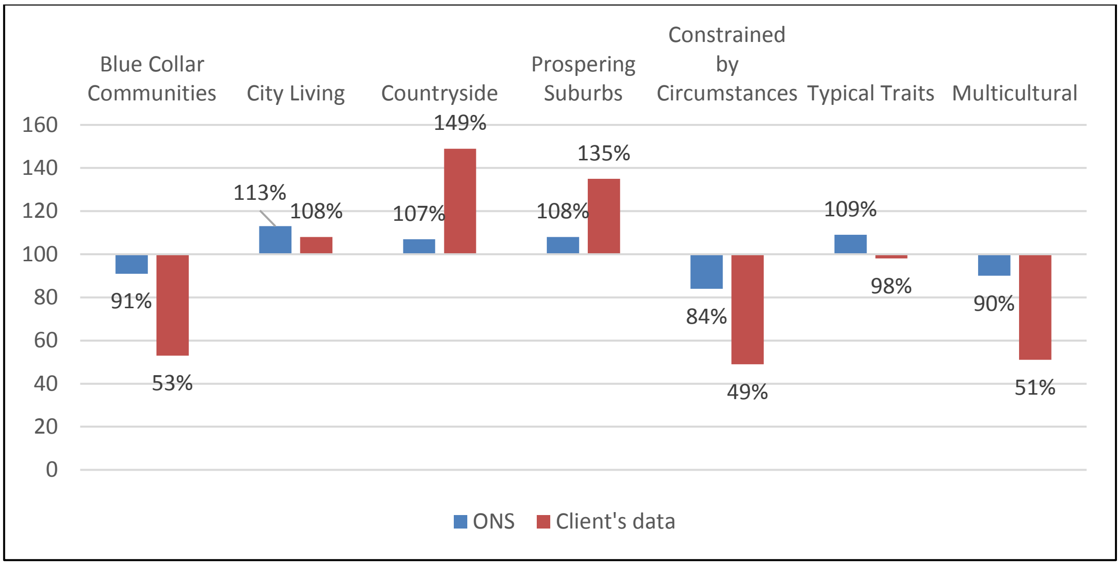


Figure 6.9. Online preferences among OAC Supergroups based on actual grocery sales data

Again, the majority of the highest on-line users (with an e-share from 11% to 37% in 9 areas) belong to the Affluent Achievers consumer classification category which incidentally have a strong preference to shop at the client's stores (with an index of 122 compared to the national UK consumer index (CACI, 2013).

As Figure 6.9 shows, there is a good deal of agreement between the ONS survey data and client's on-line customers in terms of the main groups who are active in e-commerce. Typical traits and city living provide a very good match. However, there and some interesting differences between these two data sets for Blue Collar Communities, Constrained by Circumstances and Multicultural. These latter groups are the lowest income groups in the UK population and these differences might simply be explained by the fact that that this particular supermarket chain is not very popular among customers belonging to these OAC groups. The higher rate for the Countryside group is also interesting which could be explained by low accessibility to the nearest client's supermarket of over 3500 sq ft (with an average distance of 4km).

In terms of frequency of buying on-line among OAC groups the pattern is very similar to the distribution of on-line expenditure with customers belonging to Prospering Suburbs and Countryside categories being the most regular on-line customers with an index of 131, although customers within Countryside group are spending more on their groceries.

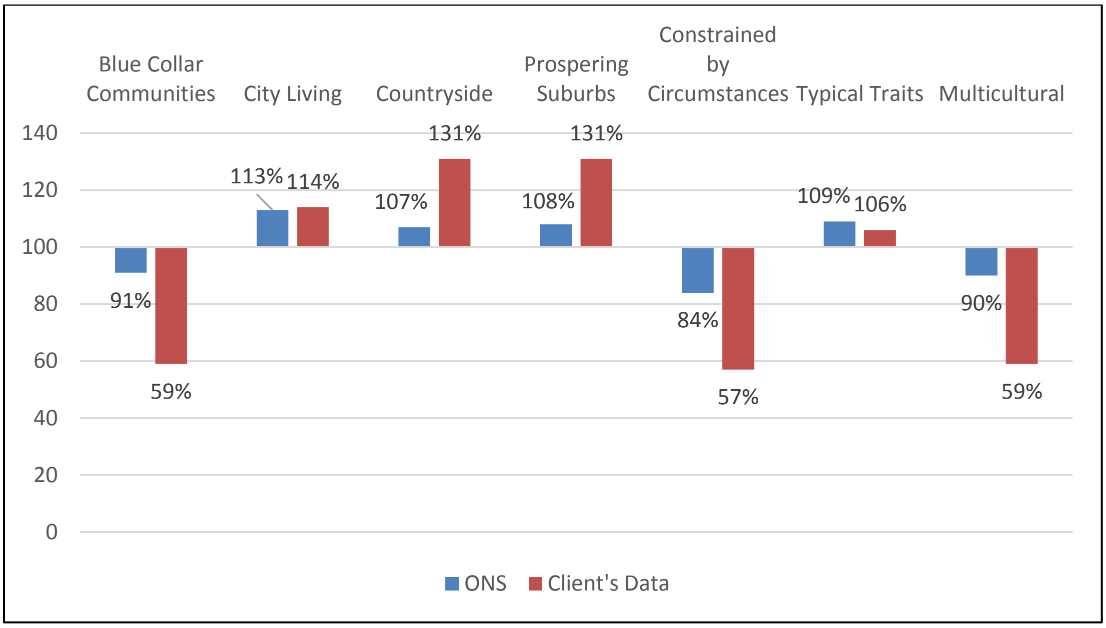


Figure 6.10. Frequency of online purchasing among OAC Supergroups based on actual grocery sales data

Analysing the distribution of actual data there is a conclusion that client's on-line customers are more likely to belong to the higher social class and buy their groceries on-line despite close proximity to physical stores. Moreover, there is a strong indication that customers living in rural areas are also more likely to employ the on-line channel to buy their groceries due to restricted accessibility to the nearest large grocery stores. Therefore, a higher demand for on-line groceries will be expected in rural areas with a density less than 2000 inhabitants per sq km and from more affluent areas with higher concentrations of customers within the higher social classes – Affluent Achievers and Rising Prosperity. The relationship of distance, store size and on-line expenditure will be explored in further chapters in more detail. The next section will analyse the on-line customer profile by product type.

6.5. Online expenditure within product categories

This section provides an analysis of actual on-line expenditure by product type and will also make comparisons to survey data obtained from CACI and national statistical data. CACI provided estimated data for online expenditure within COICOP product classification (see Appendix B). Client provided data using their own product classification for 52 categories which was aggregated into 12 major food and drink product categories. Family Food Data for 2013 from the Department for Environment, Food and Rural Affaires (DEFRA) and Mintel research data was applied to allow a comparative analysis with actual and CACI estimated data. The data from three sources (DEFRA, CACI and Supermarket) was aggregated into 9 major product categories to correspond with Mintel data.

In terms of product categories, customers' on-line expenditure reflects general grocery spending with the largest proportion (almost 50%) spent on meat and fish, fruit and vegetables, and alcoholic drinks (Figure 6.11). The least popular on-line products are 'ethnic food' followed by confectionery goods and snacks, soft drinks and bread and desserts, that when combined only make up 20% of the total online expenditure.

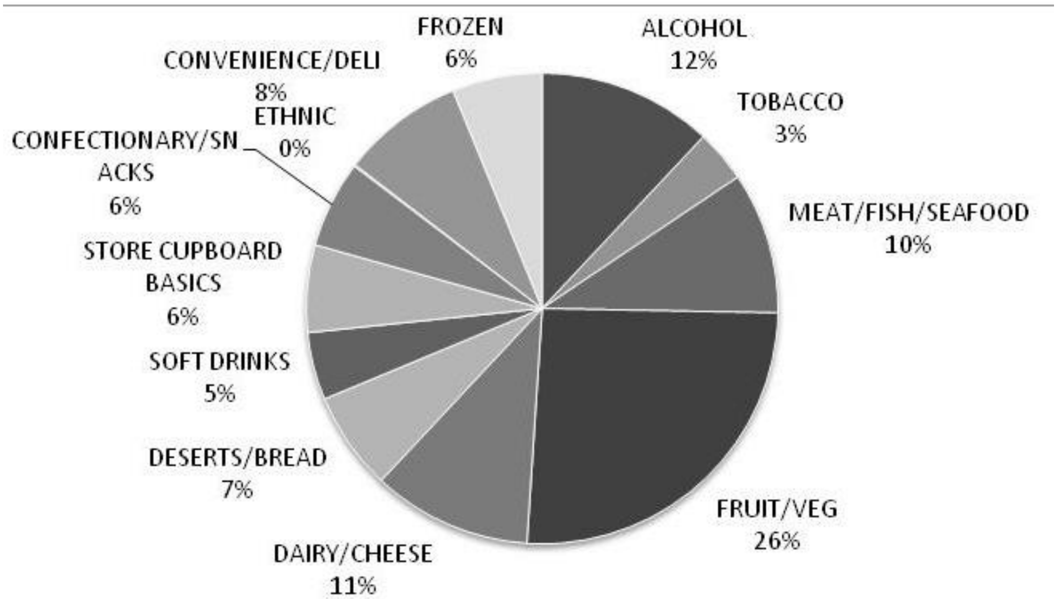
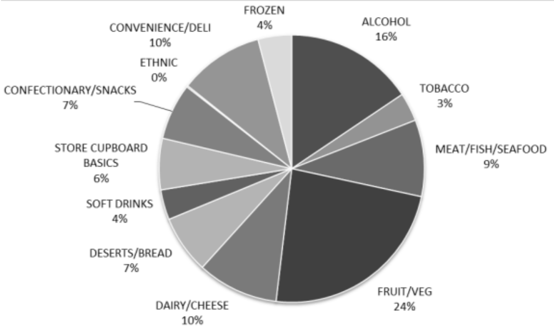


Figure 6.11 Supermarket Online Food Sales. Source: Client's Data April-June 2015

The total supermarket grocery expenditure within the major grocery product categories and the share of actual on-line expenditure in total supermarket expenditure are presented in Figure 6.12 and Figure 6.13.



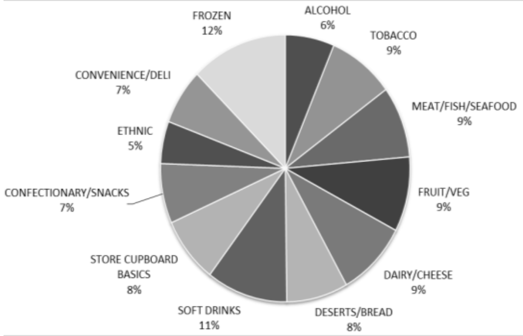


Figure 6.12. Client’s Total Food Sales

Figure 6.13. Client’s Share of Online Sales in Total Food Sales

Figure 6.12 and 6.13 clearly demonstrate very similar customers’ spending patterns in both channels – on-line and physical stores. The slight variation is within alcohol, confectionary, convenience and fruit and vegetables categories with customers spending marginally more on fruit and vegetables on-line compared to in store purchases and vice versa with alcohol and confectionary expenditure 4% and 1% greater respectively in physical stores compared to on-line channel due perhaps to an impulse buying factor when customers are enticed into purchasing by special promotions and POS displays. The convenience products are more likely to be purchased in store due to their immediate availability. Figure 6.12 shows similar distribution of client’s on-line share in total food sales within all 12 major product categories. Soft drinks and frozen products (with a higher share of 11% and 12%) are more likely to be purchased on-line as consumers prefer bulky and heavier items being delivered to their door. Moreover, on-line customers are more interested in ethnic foods with 5% share in total food expenditure compared to less than 1% in total food sales (Figure 6.13). The total supermarket chain on-line share is 9% within the major grocery categories, which is higher than national figures of 4.4% (IGD.com, 2016). This fact indicates that client’s on-line customers tend to belong to the higher social class which is supported in literature review discussed in Chapter 3.

Figure 6.14 provides further analysis of expenditure within product categories and compares the degree of on-line expenditure within nine product categories of client’s data against three reputable sources – DEFRA and the market research agencies Mintel and CACI.

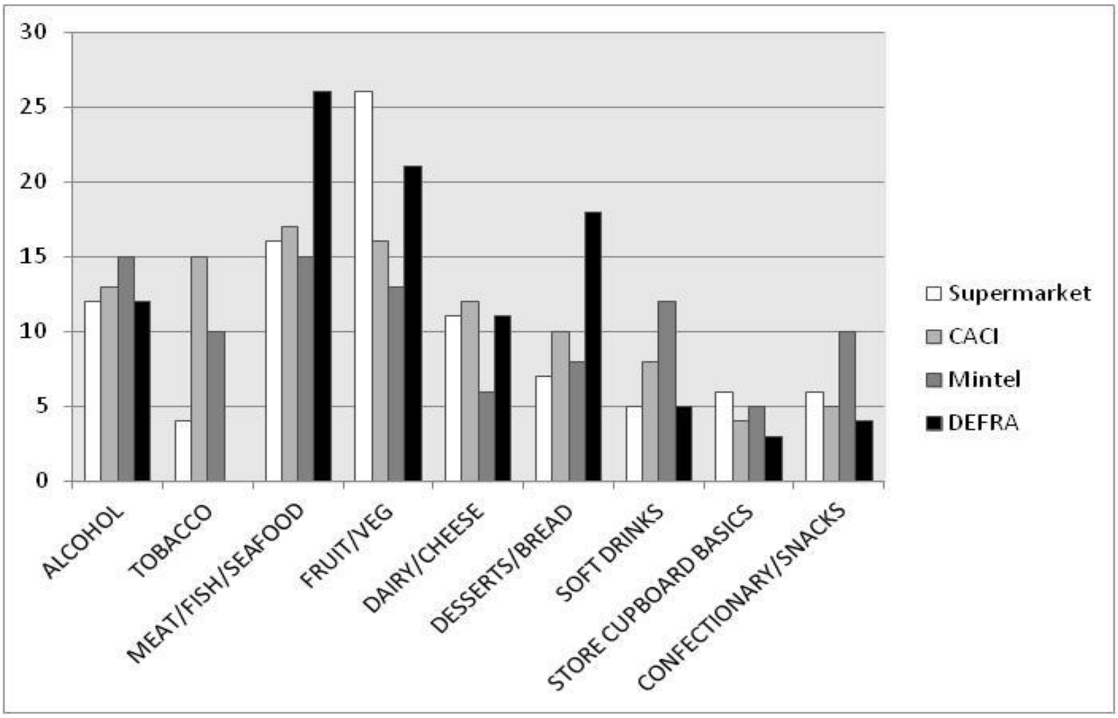


Figure 6.14. Comparative analysis of online expenditure by product categories. Source: Client's Data April-June 2015, CACI data 2013, Mintel 2013, Defra 2013

Comparing the Mintel and CACI estimated data by product category to the actual client's data there are slight variations in on-line expenditure with the higher spending on meat and fish products and alcohol and lesser expenditure on fruit and vegetables shown in the former studies. Interestingly, the estimated and actual on-line expenditure significantly varies for tobacco products with CACI's estimation of 15% in total on-line expenditure compared to only 4% of client's data. This may be largely due to the demographic profile of the client's typical shopper which belongs to the more affluent social categories, where smoking rates are distinctly below the national average (with an index of 70 compared to 100 UK base (CACI, 2013). Moreover, the fact that client's on-line shoppers buy significantly more vegetables and fruits with the quarter of the total on-line expenditure compared to 13% and 16% Mintel and CACI figures respectively, supports client's shopper demographic profile as belonging to the higher social categories.

Moreover, client's on-line shoppers spend considerably less on soft drinks (5%) compared to the Mintel data of 12%. At the same time, the client's on-line shopping basket reflects the general grocery expenditure of the average UK consumer within several categories including soft drinks (5%), alcoholic drinks (12%), dairy products

(10%) and confectionery with 4% of total grocery expenditure. The major variations in on-line and total grocery expenditure are within the meat and fish categories with over a quarter of total grocery expenditure spent on these products compared to only 10% of total on-line expenditure. The bread and desserts products also unlikely to be purchased via an on-line channel with only 7% on-line share in total grocery expenditure in this category. In general, consumers are cautious with buying on-line ethnic, perishable and impulse products.

The market share of on-line expenditure by product category is presented in Figure 6.15. The wine sector has the highest on-line share with almost 7% of the total due to customers buying heavy, bulky, long lasting goods on-line with 25% of wine bought on-line (Union Press Ltd, 2014). In 2012 the wine market was worth £800million with an annual growth rate of 21% since 2005 with the major supermarkets opening their own on-line wine stores due to the increased demand. Tobacco goods have the smallest market share due to more popular distribution channels, e.g. off-licence and petrol stations (Mintel, 2013).

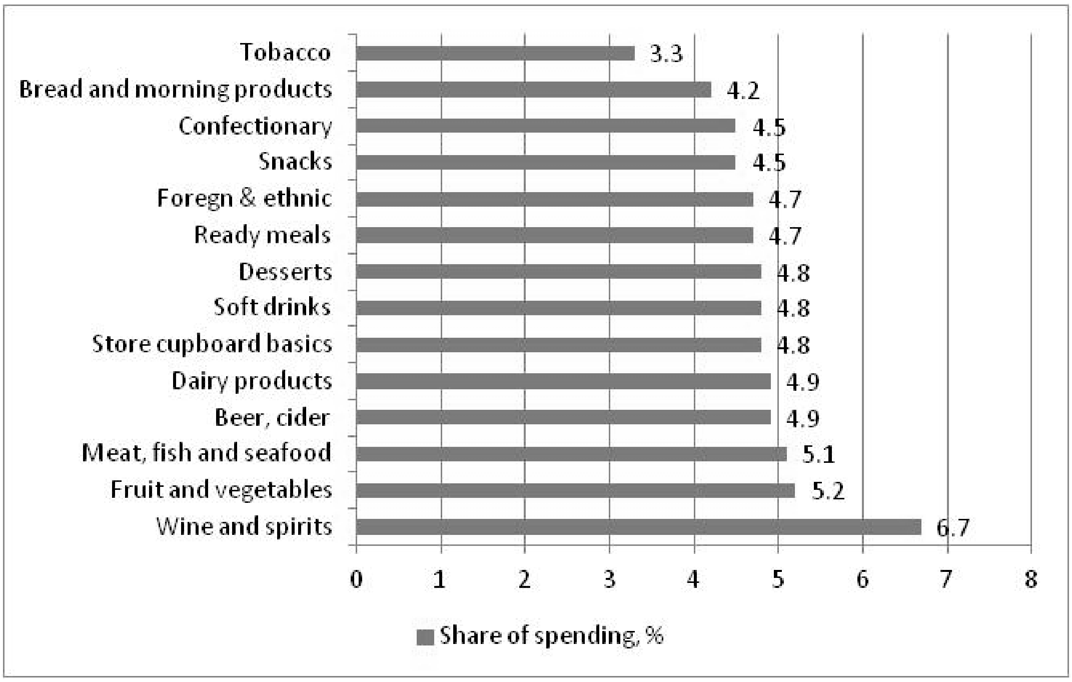


Figure 6.15. Online: Share of food and drink spending by category, 2012. Source: Mintel, 2013

The distribution of on-line usage by product category indicates that on-line grocery shoppers are more likely to be British, middle aged and belong to the higher social class.

6.6. The summary

This Chapter attempted to estimate online demand for the study area based on the actual and estimated data. Firstly, the analysis showed the strong relationship between estimated data provided by CACI and actual on-line expenditure by client's customers. Moreover, research estimates of on-line demand showed very similar distribution of predicted on-line expenditure in comparison to CACI estimates. The geodemographic analysis of online customers identified those areas with high concentration of young people aged between 25 to 44, families with two or more children and individuals belonging to the higher social classes AB and C are likely to generate the highest on-line demand in the study area. Although, analysis of actual data didn't establish the relationship between on-line expenditure and demographic characteristics, i.e. age and gender. Consequently, family composition and social class are two major demographic characteristics of on-line customers. Moreover, a higher demand for on-line groceries will be expected in rural areas with a density less than 100 inhabitants per sq km and from more affluent areas with higher concentrations of customers within the higher social classes – Affluent Achievers and Rising Prosperity. The more detailed analysis of the relationship between accessibility and on-line expenditure will be explored in Chapter 8. The next chapter will introduce the Spatial Interaction Model for estimation of residential grocery demand (total grocery demand) for the study area.

Chapter 7. Building a face to face model for the grocery market

7.1. Introduction

In this Chapter a Spatial Interaction Model (SIM) will be designed and applied to estimate the distribution of consumer spending at grocery stores across the study region (a face to face model). This is necessary before the construction of a model could be addressed which includes online expenditure and sales. The justification and details of this technique were outlined in Chapter 4. This Chapter describes the stages required in designing a SIM of estimated residential demand and its distribution across physical stores. First, section 7.2 provides an overview of the SIM and its components in terms of demand, supply and interaction. Section 7.3 outlines the demand side of the model and plots the spatial distribution of the demand estimates across the study area. The supply side of the model will be considered in section 7.4 with an overview of the grocery stores in the region and the justification for floorspace to be used as an attractiveness factor. Section 7.5 introduces the distance decay parameters of the model. Sections 7.6 and 7.7 provide an overview of model disaggregation by brand attractiveness and distance decay. Finally, section 7.8 presents the model results and evaluates the estimated data based on three variables – grocery market shares, spatial analysis of estimated and actual data and comparison of observed and predicted distance travelled.

7.2. Modelling shopping flows in the Study Area

As outlined in Chapter 4, a production-constrained entropy maximising SIM is used in this research to estimate customer flows from their home location to grocery stores. The model can be written as follows;

$$S_{ij} = A_i O_i W_j exp^{(-\beta c_{ij})} \quad (7.1)$$

Three components are required to build the SIM and to allocate residential grocery expenditure in the study area. The first component is available weekly grocery expenditure in each postal sector (O_i). Supply data onthe attractiveness of each grocery store (W_j) is the second part of the SIM. Finally, the distance between residential areas (i) and grocery stores (j) is required to measure accessibility or relative 'cost' of distance travelled (C_{ij}) which has an associated distance decaying parameter (β), which controls for ease of travel. Initially, the model was run based on these 3 measures: demand data, attractiveness of the store in terms of each stores' grocery floorspaceand the distance between customer location (postal sector) and store destination. The next runs of the model include disaggregation by the distance decay parameter (β), household types (*m)*, and brand attractiveness among five Acorn geodemographic categories (α^m). Formula (7.2) shows the model form when disaggregated:

$$S_{ij}^m = O_i^m A_i^m W_j^{\alpha^m} \exp(-\beta^m C_{ij}) \qquad (7.2)$$

A MS Excel spreadsheet programme was used to run the model with the application of statistical programmes (Minitab and SPSS) to test the model's accuracy and validity. The model calibration process was applied to find the most appropriate parameters to obtain the best fit between estimated and actual data. Average Distance Trip (ADT) is an appropriate variable to evaluate the model. This uses a negative exponential function for trip distance. The ATD formula, which estimates the difference between predicted and observed ATD can be described as follows:

$$ATD = \frac{ATD^{\Pr ed}}{ATD^{Obs}} \qquad (7.3)$$

Where predicted and observed ATD are the sum of customer flows (predicted (S)

and estimated *($\hat{S}$)*

$$ATD^{\Pr ed} = \frac{\sum ijSijCij}{\sum ijSij} \qquad (7.4)$$

$$ATD^{Obs} = \frac{\sum ij\hat{S}ijCij}{\sum ijSij} \quad (7.5)$$

Batty and Mackie (1972) state that this is the most appropriate calibration statistic to use for a SIM due to the fact that if the model can replicate the average travelled distance or 'cost' of distance then it is likely to estimate grocery sales more effectively.

The data required for calibration has been derived from the individual client's customer transactions which were aggregated to the Output Area and Postal Sector levels of geography. Client's in house location team, based on the customer flows from Output areas to the stores, produced revenue estimations for each store. Based on customer home postcode and distance to client's stores (excluding any distances of over 100km) the following ATD were used to calibrate the model: supermarkets - 2.2 and convenience stores – 1.0 (Median for 2013).

Effective model calibration depends on the quality of the observed data. The detailed analysis of the observed (Nectar card) data was outlined in Chapter 5. Some model underestimation is expected due to the fact that not all client's customers participate in the loyalty scheme. To eliminate this problem the final model was also calibrated against store revenues (derived from POS data) for the same period of time. A comparative analysis of estimated and national grocery market shares was additionally used to test model's validity. Each of these stages of model design are outlined in the following sections in more detail.

7.3. Demand

The demand estimation procedure for residential grocery expenditure was outlined in Chapters 4 and 6. This section will describe the grocery demand estimation in more detail. The following formula represents grocery residential demand, segmented by different household types:

$$O_i^m = e^m n_i^m \qquad (7.6)$$

Where:

O_i^m is the total grocery expenditure available in zone (i) by consumer type (m)

e^m is the value of average weekly expenditure by consumer type (m)

n_i^m is the number of consumers by type (m) in zone (i)

To calculate residential demand two types of data were used to estimate household expenditure within the study area. First, is the number of households in each of the 791 postal sectors. In 2013, there were 2.2million households in Yorkshire and Humberside (ONS, 2013). The data on households was provided by CACI. Secondly, the number of households (2.2million) was multiplied by the average household weekly expenditure data taken from the LCFS data for 2012. The LCFS is based on a sample of 5000 households who completed the survey during a two week period. The LCFS uses the ONS OAC derived from the census data at the Output Area level which was described in Chapter 3. The average household expenditure data is based on the classification COICOP (Classification of Individual Consumption by Purpose) which consists of 12 major categories (described in Chapter 6). In 2013 the average weekly grocery household expenditure for the study area was £67, which included expenditure on food, non-alcoholic drinks, alcoholic drinks and tobacco (see Table 6.1). To reflect the differences between different population groups, the demand has been disaggregated by grocery expenditure based on the demographic output area classification (OAC). Table 7.1 outlines the average household expenditure by OAC groups.

Table 7.1. Average household expenditure by OAC groups

OAC Group	Name	Food & non-alcoholic drinks (£)	Alcohol and tobacco (£)	Total (£)
1A	Terraced blue collar	47.4	13.4	60.8
1B	Younger blue collar	51	13.3	64.3
1C	Older blue collar	50.1	11.7	61.8
2A	Transient communities	42.5	11.8	54.3
2B	Settled in the city	51.2	10.1	61.3
3A	Village life	61.6	13.3	74.9
3B	Agricultural	67	12.5	79.5
3C	Accessible countryside	62	12.7	74.7
4A	Prospering younger families	61.6	12.7	74.3
4B	Prospering older families	64.2	13.9	78.1
4C	Prospering semis	58.5	11.2	69.7
4D	Thriving suburbs	64	14.7	78.7
5A	Senior communities	38.4	10.3	48.7
5B	Older workers	44.7	10.3	55
5C	Public housing	42.3	16.1	58.4
6A	Settled households	53.9	11.4	65.3
6B	Least divergent	58.2	11.5	69.7
6C	Young families in terraced homes	48.6	11.2	59.8
6D	Aspiring households	56.7	11.6	68.3
7A	Asian communities	56.8	11	67.8
7B	Afro-Caribbean communities	49.6	8.3	57.9
	Average Total	**54.8**	**12**	**66.8**

Source: ONS Family Spending 2011 (2012)

Table 7.1 shows that areas with high concentrations of Prospering Suburbs and Countryside communities are likely to have the highest grocery demand with average weekly grocery expenditures of £74.2 and £76.4 respectively. The total grocery estimated demand for the study area is £148million which corresponds with CACI estimated data of £149 million. Figures 7.1 and 7.2, respectively, show the distribution of CACI's estimated data compared to demand estimated by the author herself in the study area.

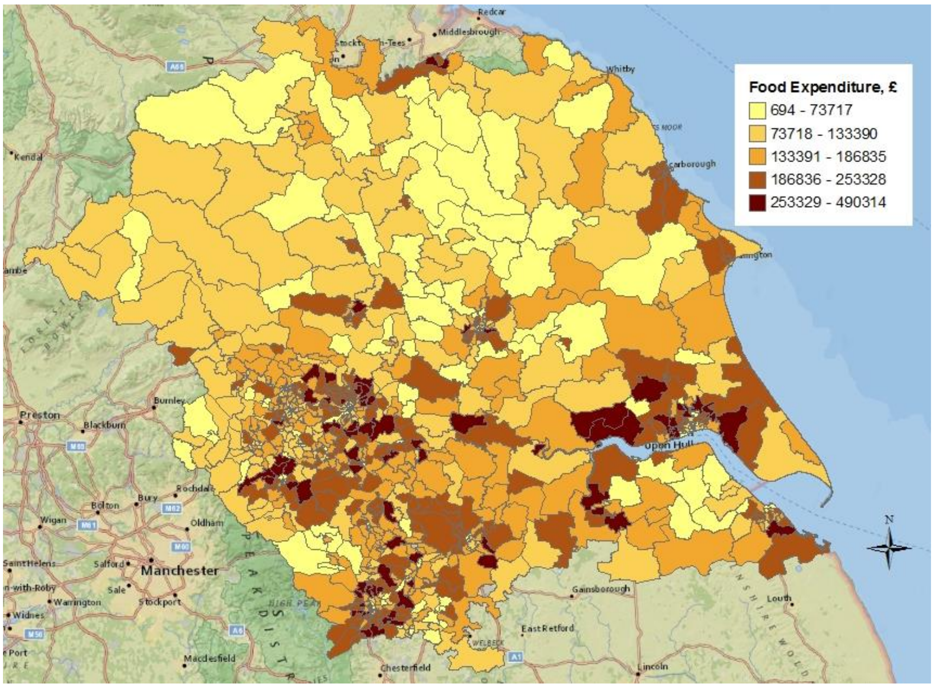


Figure 7.1. CACI weekly reseidential grocery expenditure

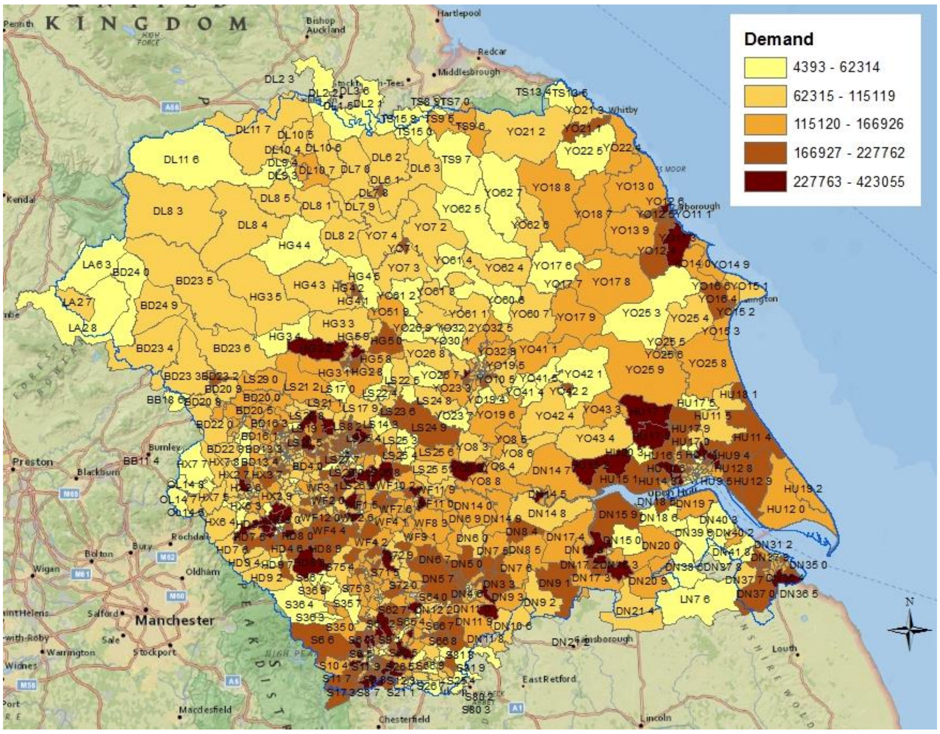


Figure 7.2. Weekly demand estimates in Yorkshire and Humberside by author

The maps show the estimated demand distribution across Yorkshire and Humberside (in £ per week) with the highest demand unsurprisingly in urban areas due to high population density and high concentrations of Prospering Suburbs communities in some locations. Note how similar the spatial distributions are in both maps.

7.4. Supply

The supply side of the SIM relates to the available grocery stores in the region, and their attributes. The attractiveness of the retail unit in SIMs is normally associated with its size, as larger stores can offer a wider selection of products and services, have better provision in terms of parking spaces and normally offer lower prices (Birkin et al, 2002). In this research grocery store floorspace is used as the main attractiveness factor. The floorspace data was supplied by client for 1812 grocery stores in the study area with sizes ranging from 400sq ft to almost 116,000 sq ft, with an average grocery floorspace across all postal sectors of 6300 sq ft. Table 7.2 shows the distribution of grocery floorspace among the major different grocery retailers in the region.

Table 7.2. The grocery retailers in Yorkshire and Humberside

	Floorspace (%)	Stores
Aldi	4	47
Asda	13	73
Co-op	12	355
Iceland	2	40
Lidl	5	47
Morrisons	18	71
Tesco	18	167
Waitrose	1	7
Spar	2	143
Sainsbury's	9	131
Costcutter	2	154
M&S	5	48
Londis	2	130
Premier	3	214
Nisa	1	61
One Stop	1	74
Others	3	50
Total	**100**	**1812**

Table 7.2 shows that Morrisons and Tesco supermarket chains have the largest presence, with 18% each of the total available grocery floorspace in the area. Although Morrisons has a lower national market share than Tesco, it is not surprising that their regional grocery market share is higher due to its Yorkshire origins, with its headquarters in Bradford and a high concentration of stores in the study area. They are followed by Asda and Co-op grocery retailers with 13% and 12% respectively of the total grocery floorspace in the study area. The Co-operative group also has the highest number of stores (355) across the study area. At the time of the research German based discounters Aldi and Lidl had a relatively small combined share of 9% in the total grocery floorspace. In the last few years these retailers have significantly increased their presence in the area. High end supermarket chain Waitrose has the smallest share of floorspace of only one percent, although it too has plans to try and develop in the north of England. Others retailers include high end supermarket chain Booths (with three stores in affluent areas of West Yorkshire), small local supermarket chains (Budgens, Ramsdens, Booths and Proudfoot) and low cost retailers Heron Foods and Farmfoods. Figure 7.3 shows the distribution of grocery stores in the study area in terms of their size.

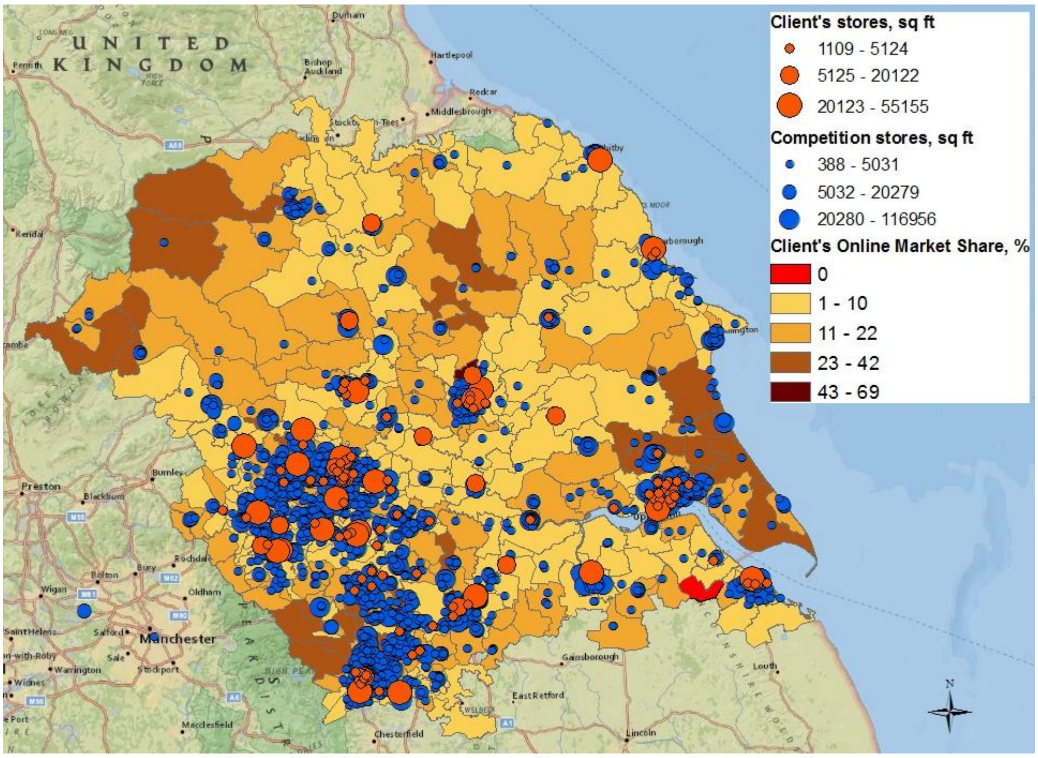


Figure 7.3. Distribution of the grocery retailers in Yorkshire and Humberside

Figure 7.3 shows that the highest concentration of grocery retailers is unsurprisingly around the large cities of Leeds, Bradford, Kingston-upon-Hull and Sheffield. The rural areas have a more limited availability of grocery stores with no presence of grocery stores in some postal sectors or with accessibility to only one small grocery store of less than 5000sq ft.

7.5. Distance

The straight line distances between origins *i (*postal sectors) to destinations *j* (grocery stores) were calculated based on the X and Y coordinates (centroids) for each postal sector within the study area. These coordinates, which are based on the British National Grid system, were obtained from EDINA – UKBORDERS[3] data set (UK Data Service, 2012). The coordinates for the grocery stores (1812 in total) were provided by partner organisation. To calculate the straight line distance (d) between origin zone (i) and shopping destination (j) the following formula was applied:

$$d_{ij} = \frac{\sqrt{((x_j - x_i)^2 + (y_j - y_i)^2)}}{1000} \tag{7.7}$$

The average minimum distance from a customer location (postal sector) to a store destination across the study area is 1.39 km. The study area has an extensive road networks and the rural and urban areas are well connected. In these circumstances application of the straight line distance technique would be adequate.

7.6. Disaggregation by brand attractiveness

Many commentators have noted that in a competitive environment such as grocery retailing, brand attractiveness is very important in helping retailers to maintain or

[3]Census and Digital Boundary Data provided by UK Data Service which is copyright of the Crown

increase their market share (Birkin et al, 2010). To sustain customer loyalty, retailers use various loyalty schemes (special offers, vouchers, reward point schemes etc.). In addition, grocery retailers have built that brand attractiveness around certain key target markets. For example, Waitrose and Sainsbury's supermarket chains will not perform as well in low income areas as consumers in these areas tend to prefer less expensive chains such as ASDA or one of the discounters like Aldi or Lidl. Within the SIM, brand loyalty can be captured through the alpha (α) value, with disaggregation by brand attractiveness among consumer types *(m)*. In SIM the parameter alpha is the power function associated with floorspace, $W_j^{\alpha^m}$, making floorspace more or less attractive to different types of consumers (by Acorn category for example) (see figure 7.2). In this research the alpha parameters were derived from the work of Thompson et al (2012) and Newing (2013). They created brand location quotients based on Axciom's research opinion poll in combination with the OAC to identify preferences towards ten major grocery brands (see Table 7.3)

Table 7.3. Brand location quotients for use in disaggregated SIM

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: Thompson et al, 2012

The table shows that the quotients are all scaled around a value of one (to five decimal places) due to the high sensitivity of the model as a result of changes to this power function. The brand location quotients by OAC groups were converted into attractiveness values among five different Acorn demographic categories using Pivot Table technique in Excel (see Table 7.4). The population for each postal sector has been divided within associated output area classifications and merged by Acorn classifications.

Table 7.4. Brand attractiveness values among Acorn demographic categories

Acorn	**1**	**2**	**3**	**4**	**5**
Aldi	1.00046	0.99967	1.00027	0.99936	0.99941
Asda	0.99783	0.99880	0.99923	1.00085	1.00199
Booths	1.00043	1.00043	1.00043	1.00043	1.00043
Budgens	1.00043	1.00043	1.00043	1.00043	1.00043
Co-Op	1.00047	0.99917	1.00072	1.00058	0.99910
Costcutter	1.00043	1.00043	1.00043	1.00043	1.00043
Farmfoods	1.00043	1.00043	1.00043	1.00043	1.00043
Heron Frozen Foods	1.00043	1.00043	1.00043	1.00043	1.00043
Iceland	1.00043	1.00043	1.00043	1.00043	1.00043
Sainsbury's	1.07245	1.040351	1.03303	0.989	0.986564
Lidl	0.99986	0.99980	1.00014	1.00069	1.00019
Londis	1.00043	1.00043	1.00043	1.00043	1.00043
M&S	1.00208	1.00458	0.99996	0.99915	0.99856
Morrisons	0.99966	0.99945	1.00103	1.00002	1.00028
Nisa	1.00043	1.00043	1.00043	1.00043	1.00043
One Stop	1.00043	1.00043	1.00043	1.00043	1.00043
Planet Organic	1.00524	1.01045	1.00103	0.99902	0.99656
Premier	1.00043	1.00043	1.00043	1.00043	1.00043
Proudfoot	1.00104	0.99979	1.00045	0.99939	0.99880
Ramsdens	1.00043	1.00043	1.00043	1.00043	1.00043
Spar	1.00043	1.00043	1.00043	1.00043	1.00043
Tesco	1.00104	0.99979	1.00045	0.99996	0.99880
Waitrose	1.00524	1.01045	1.00040	0.99902	0.99656

7.7. Disaggregation by beta value

In combination with distance described in section 7.4, different beta values can be used to reflect the desire of some customers to travel further to their chosen store due, for example, to car ownership which makes accessibility to the store relatively easier. Consumers with higher disposable incomes may also travel longer distances to their favourite shop (Birkin et al 2010). In this research the βvalue is allowed to vary by location with the assumption that customers living in more rural areas, with more limited accessibility, are willing to travel longer distances in comparison to city dwellers who have a variety of shops in close proximity. To classify customer locations from 'mostly rural' to 'mostly urban', population density values have been applied. The average population density is 2110 people per sq km per postal sector. The acceptable practice is to consider areas below average population density as being rural and consequently, areas with density above average value are classed as urban. In this research four β values were used in the model to reflect more detailed customer locations which were divided by equal breaks (Table 7.5)

Table 7.5. Beta values applied in model

Population density, per sq km	Beta value
8.24 – 3787.85	0.3
3787.85-7575.70	0.4
7570.70-11363.56	0.5
11363.56-15151.41	0.6

The initial beta value of 0.4 was identified as the best fit for the model in terms of correlation between estimated and actual grocery revenues.

7.8. SIM results

The aim of face to face SIMs is to replicate the observed data with a high degree of accuracy. In this research three methods were applied to validate the data. First, is the estimation of retailer's market shares in the region in comparison to their actual market shares (as estimated by floorspace). Secondly, spatial analysis of estimated and actual client's grocery revenue was undertaken. Finally, actual client's stores (131 in total) revenue per sq. ft. was analysed in comparison to the estimated. Table 7.6 summarises the outcomes of the model in terms of grocery retailer market shares and estimated weekly revenues.

Table 7.6. SIM results – Grocery Retailers Market Shares

Name	National Market Share (%)	Predicted market share (%)	Weekly revenue, £	Weekly revenue, per sqft, £
ALDI	5	4	3367088.24	10.2
ASDA	16	14	16813978.40	11.3
ICELAND	2	2	311322.26	6.3
BUDGENS			145629.27	10.5
CO-OPERATIVE	6	10	13646276.85	10.4
LIDL	4	5	4955168.72	11.0
M&S			13613666.85	11.3
MORRISONS	11	15	22606869.80	10.9
ONE STOP			854673.36	10.8
SAINSBURY'S	17	12	15883344.65	10.6
SPAR		2	1996541.90	9.7
TESCO	28	20	25363158.45	10.4
WAITROSE	5	2	1941292.96	11.3

Table 7.6 shows that Tesco has the largest market share, although it is lower in comparison to the national data. In contrast, Morrisons has higher predicted market share in comparison to national figures. Although, in comparison to the regional

market shares based on the grocery floorspace (see Table 7.2) Morrisons and Tesco predicted market shares have closer fit of 15% and 20% respectively. Sainsbury's predicted market share of 12% is also below national data of 17% but has a closer fit with regional market share of 9%. Interestingly, Co-op with a low market share nationally (6%), have a large presence in the study area with a predicted market share of 10% which corresponds with the regional share of 12%. Overall, the market shares results showed a close fit with regional data in particular.

In terms of weekly revenues the most successful supermarkets seem to be Asda, Waitrose and M&S with a weekly revenue of £11.3 per sq ft. Iceland and Spar are estimated to be the least successful retailers with a revenue of less than £10 per sq ft.

Figure 7.7 shows the leading retailer (by market share) in each postal sector across the study area.

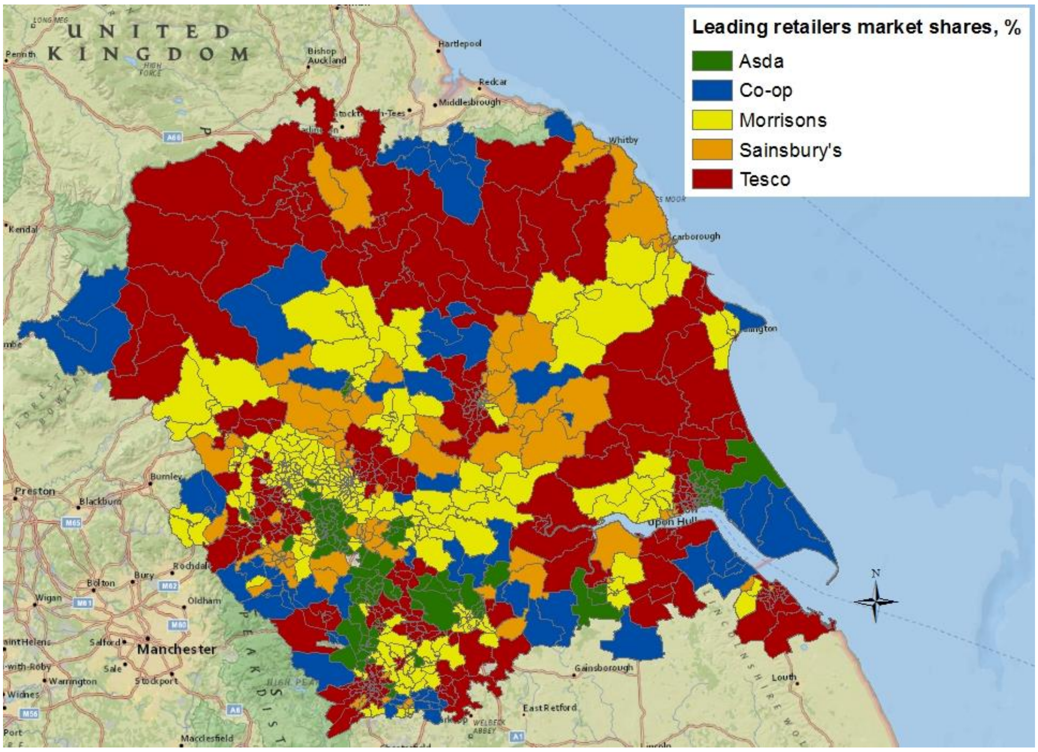


Figure 7.7. Leading Retailers Market Shares in the study area

Tesco dominates the grocery market being the leader in the majority of postal sectors. Morrisons has a dominance in the city of Leeds and in the north of the study area. Sainsbury's have a large presence in more affluent areas of York and Leeds.

Figures 7.8 and 7.9 shows the spatial distribution of client's estimated and actual grocery revenue.

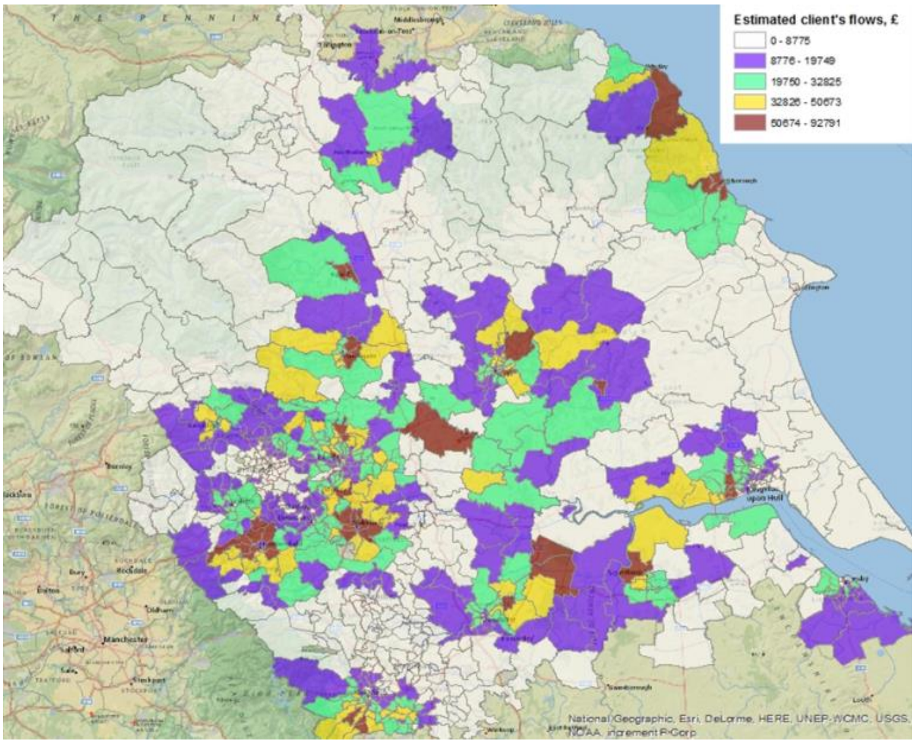


Figure 7.8. Client's estimated revenue by postal sector

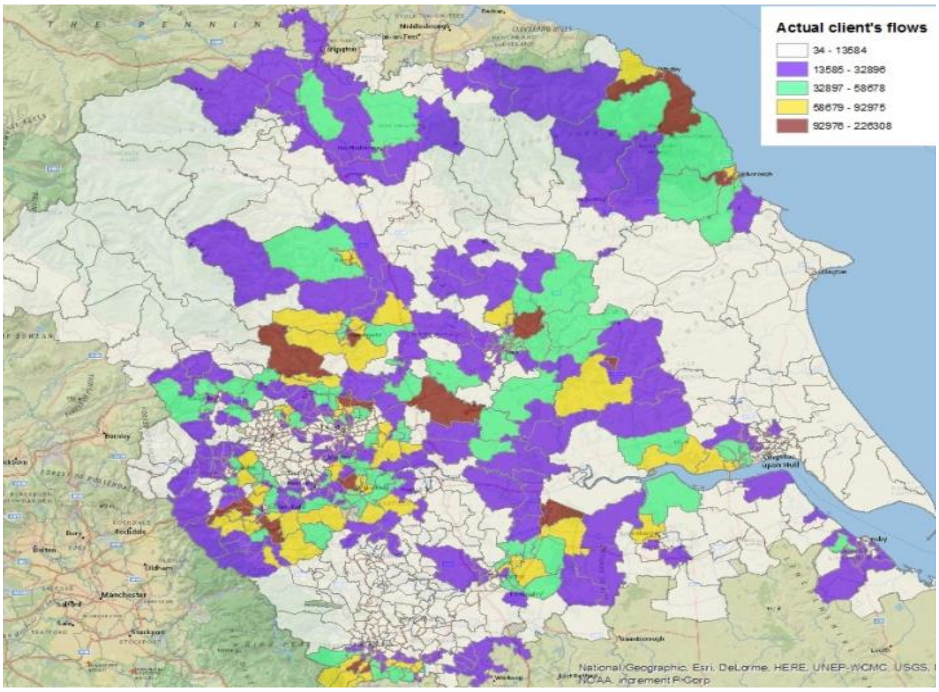


Figure 7.9. Client's Actual residential expenditure (loyalty card users only)

These maps demonstrate similar patterns of grocery expenditure/revenue between estimated and actual data, with a data correlation of 83%. Figure 7.10 demonstrates the correlation of client's actual weekly store revenues (POS) and predicted revenue across the 131 stores. The correlation between the data is 97%, although, the predictions for convenience stores are much lower in comparison to supermarkets due to the more complicated nature of the convenience market and the difficulties in identifying the nature of the catchment areas for certain convenience stores – especially those in the city centre where demand is largely driven by workplace consumers (Hood 2016).

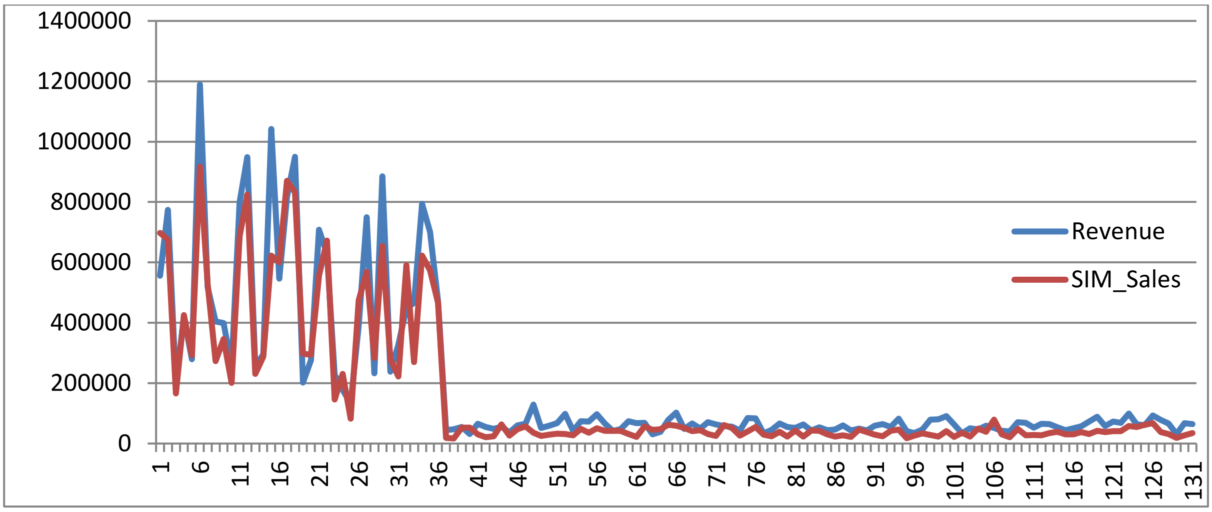


Figure 7.10. Client's stores expenditure and estimated data

The last test is the comparison of estimated and actual ATD distances. Table 7.7 demonstrates the following average results for supermarkets and convenience stores produced by the model

Table 7.7. ATD Distances produced by model

	Observed ATD	Estimated ATD
Supermarkets	2.2	2.5
Convenience stores	1.0	1.3

Observed and estimated travelled distances shows a close fit.

Figure 7.11 shows the distribution of client’s market share across the study area as estimated by the model.

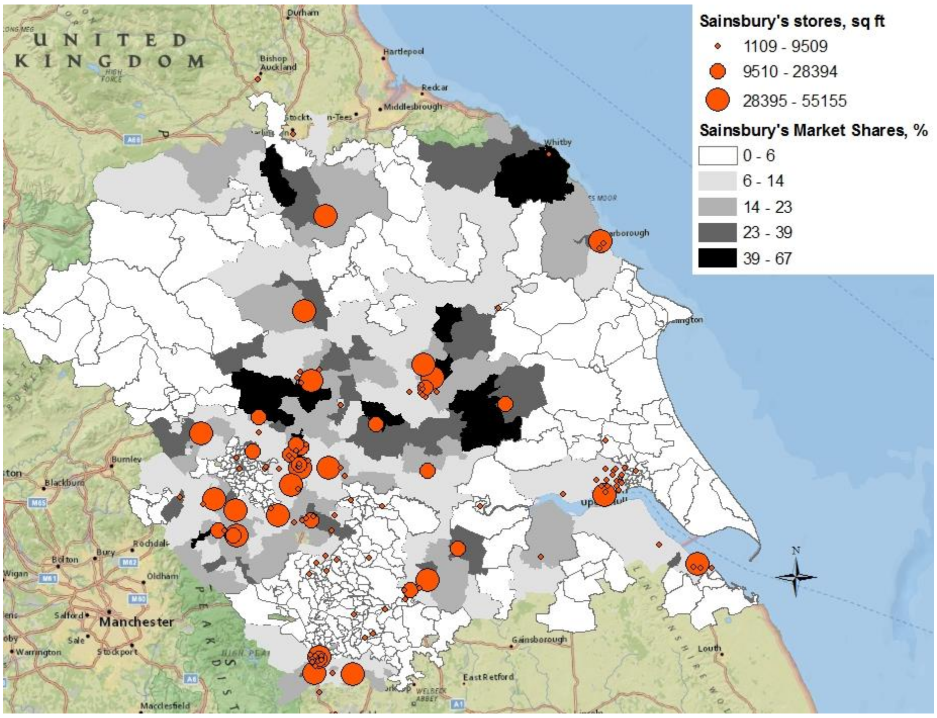


Figure 7.11. Distribution of client’s market shares in Yorkshire and Humberside

The map shows that the highest market share (67%) is estimated in some of the suburban areas of Leeds, York and in the north of the region. The lowest market share is expected in the rural areas of Yorkshire and Humberside due to low presence of the retailer in these areas and higher concentration of competitors’ stores (see Figure 7.11).

7.9. Conclusions

The aim of this chapter has been to present a face to face model which can estimate grocery sales with high accuracy. The classic production-constrained SIM was applied to estimate revenue for face to face grocery retailing grocery in Yorkshire and Humberside. The results presented in section 7.8 demonstrate that this has been achieved with a high degree of accuracy. The actual data provided by client across all 131 of their stores in the region assisted in model calibration. The model showed an 83% correlation with actual data across all 791 postal sectors and 97% correlation between observed stores sales data and results produced by the models across 131 client's stores for the same period of time. The model showed an underestimation for the convenience stores due to the more varied nature of their catchment areas, with many being driven largely by a workforce location (Hood, 2016). The model showed high accuracy in revenue estimation for supermarkets which are more likely to be visited by customers residing within the catchment area. The model has been disaggregated to reflect brand attractiveness and customer locations against store accessibility. The inclusion of alpha and beta values allowed both supply and demand to be disaggregated separately with the connection between them maintained through the re-use of consumer type data (Acorn classification) on both the demand and supply side. The model showed the predicted spatial distribution of grocery retailer's market shares with Tesco being a leader with an estimated 20% which compares to its national rate of 28% and regional estimates of 18% (based on available grocery floorspace in the region). Client's estimated market share is 12% which is lower compared to its national figure of 17% but corresponds with its regional market share of 9% of floorspace. The supermarket is expected to have higher market shares of up to 67% in the more affluent suburban areas of Leeds and York and lower market shares in the rural areas of Yorkshire and Humberside due to the low presence of physical stores in these areas of the study area.

The next chapter produces a model which attempts to incorporate another dimension of modern grocery shopping –the online channel. To introduce a new layer of online expenditure a good fitting face to face model is required which was achieved in this chapter.

Chapter 8. Exploring and modelling the spatial distribution of online users

8.1. Introduction

In chapter 7 a traditional spatial interaction model for face-to-face grocery retailing was built and calibrated, in part using client's data. This chapter builds on previous chapters relating to e-commerce and particularly the discussion relating to the major drivers of online expenditure, identified in terms of geodemographic characteristics and accessibility. This Chapter attempts to design a spatial interaction model (SIM) which includes on-line expenditure also. First, section 8.2 introduces the quadrant analysis technique which explores the relationship between actual on-line expenditure (using client's loyalty card data), customer locations and physical store provision in more detail. Furthermore, based on the results of quadrant analysis and information from previous chapters, section 8.3 provides a detailed overview of the building of the SIM which includes online expenditure. Section 8.4 outlines the model's calibration and its results. Finally, some future developments are proposed in Section 8.5 for further improvements to the model.

8.2. Quadrant analysis

To explore the more complex potential relationships between on-line buying, store provision, geodemographics and population density the quadrant analysis technique is applied which identifies strengths, weaknesses and differences of two variables and produces four types of outcomes which falls in one of four squares (Startupfactory, 2014). The following four key variables are examined using a form of quadrant analysis.

1. C prov - grocery floorspace provision by competitors stores across the study area.
2. S prov – client's grocery floorspace provision.

3. Share – e-business share in total grocery expenditure based on the loyalty scheme data provided by the client.
4. Urban – rurality of the area, or population density, calculated as number of people per square kilometre at the individual postal sector level.

For a more comprehensive analysis a ranking method has been applied with all indicators being graded from 1 to 791 starting with 1 as the highest value postal sector (the average value is therefore 396 out of 791 postal sectors). The average values for the four key indicators are presented in Table 8.1. The values lower or greater than average are considered to have higher or lower grocery stores provision in the area.

Table 8.1. Average values for the quadrant analysis indicators around Yorkshire and Humberside

Indicator	Average value
S prov	0.47 sq. feet
C prov	5.2 sq. feet
Share	10.3%
Urban-rural	0.002

Building on the geodemographic analysis in Chapter 3 the first issue is to examine concerns the interaction between client's provision of stores and urbanisation or population density and e-share. Table 8.2 demonstrates that given a lower physical channel provision in the more rural areas, on-line usage uptake is the greatest at 11.2% (with 427 observations), compared to 8.1% e-share in the more urbanised areas that have a greater presence of client's stores. Interestingly, e-share is similar in the less urbanised areas with a greater client's' presence and vice versa.

Table 8.2. Client's store provision and population density in relation to e-share

	S Prov <0.47	S prov >0.47
Urban<0.002	Average share 11.2% [N=427]	Average share 9.6% [N=44]
Urban>0.002	Average share 9.4% [N=255]	Average share 8.1% [N=65]

The second test to apply to the client's data is on the relationship between e-business share and client's provision using a ranking technique. In the 80 postal sectors with the highest client's provision the average rank of "Share" is 520, which is above the average of 396. There is a strong suggestion that there is a substitution between physical and virtual channels taking place here: with higher store provision so on-line grocery spend decreases. This substitution phenomenon (referring to when on-line purchases completely replace a trip to a physical store) has been argued elsewhere. For example, Dixon and Marston (2002) identified that 28% of their sample of 450 UK consumers in a town in southeast UK had replaced an in-store purchase. Figure 8.1 provides a visual demonstration of this effect with the horizontal axis value set to 10.3 (the average percentage value for on-line expenditure) and the vertical axis is set to the average value of 0.47 for client's store provision in the study area. The logarithmic values for store provision are used due to the small values compared to the online expenditure. The majority of instances (postal sectors) are located in the top left-hand corner, representing low on-line expenditure and higher physical store presence in the area. There is thus evidence again of a substitution effect between physical and virtual channels.

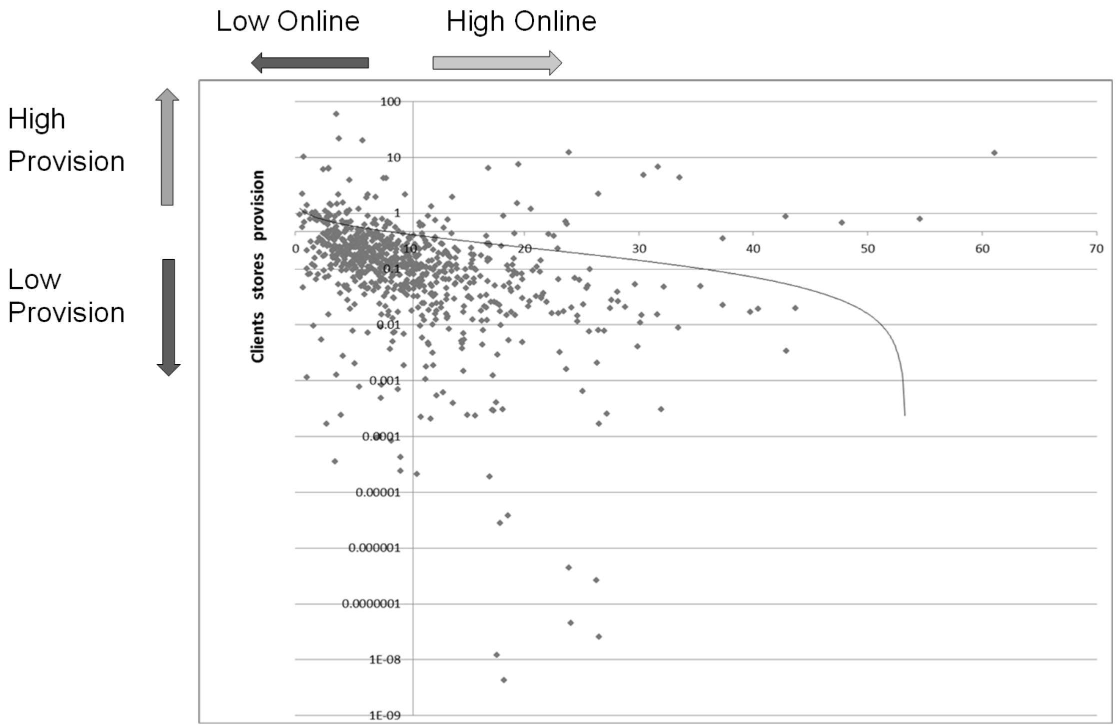


Figure 8.1. Online expenditure versus client's stores provision

Interestingly, however, in recent projects with two major UK high street retailers, CACI has established that their on-line sales actually increased as a result of greater store presence due to the effect of brand awareness and the existence of a "click and collect" service (Langston, 2011). Thus the relationship between store presence and Internet sales may vary by product type and type of location (high street versus out-of-town etc.).

The third test is to examine client's provision against competitors' provision. The hypothesis here would be that for given levels of client's provision (S prov) then low levels of Competitor provision (C prov) will tend to encourage higher levels of on-line use because there are no alternatives. The evidence seems partly to support this idea with a higher than average rank share in the areas with less client's provision (the e-shares of between 10% and 14%) and with the lowest e-share of 4% in the areas with higher client's store provision (Table 8.3).

Table 8.3. Client's and competitors store provision in relation to e-share

	C prov < 5.2	C prov >5.2
S prov <0.47	Average Rank 378 (share 10%) [N = 667]	Average Rank 262 (share 14%) [N = 15]
S prov >0.47	Average Rank 632 (share 4%) [N = 61]	Average Rank 386 (share 14%) [N = 48]

The high e-share in the areas with lower client's store presence but higher competitor store provision may indicate that these consumers favour this particular supermarket despite having a good accessibility to competitor grocery stores. These might be especially loyal consumers to the client's brand.

The instances of the rather complex relationship of on-line share, physical store provision and accessibility can be seen in the small-area geographies of client's e-usage in the maps below. Figure 8.2 shows the distribution of client's on-line market share with the additional layer of total grocery floorspace per postal sector for all 1812 grocery stores across the region. Predictably the highest concentration of grocery stores is around large cities, e.g. areas G, H and F shows the greater distribution of grocery stores in the major regional cities - Leeds, Bradford, Sheffield and York. Rural areas in the north of the region (areas A and B) have poor local grocery provision and higher on-line shares, up to 34% in some postal sectors. The opposite situation is observed in areas D and E where e-commerce activity is very low perhaps due to the large physical store presence. Area C seems to be a major anomaly. Here, a high e-commerce usage and a high level of store provision (with on-line shares of up to 34% and high grocery floorspace provision of up to 68000sq ft.) can be observed.

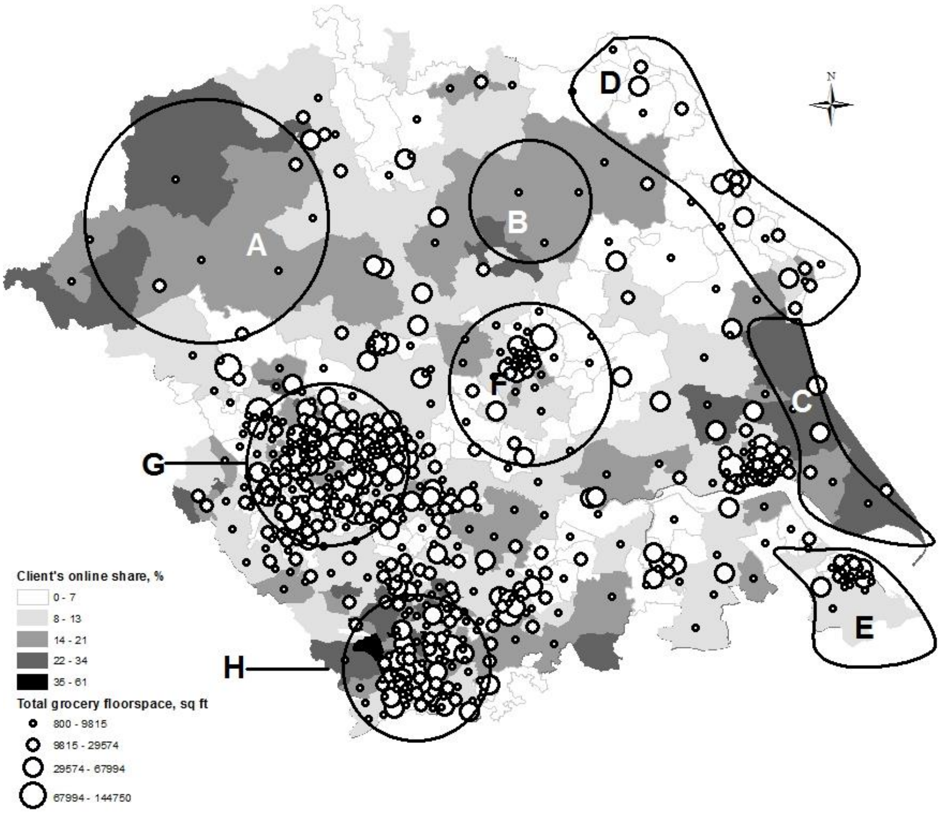


Figure 8.2. Total Floorspace Yorkshire and Humberside in comparison with online share

The Figure 8.3 demonstrates the distribution of client's online share with the proximity of the grocery stores with enlarged map of the major city in the study area – Leeds and the surrounding area

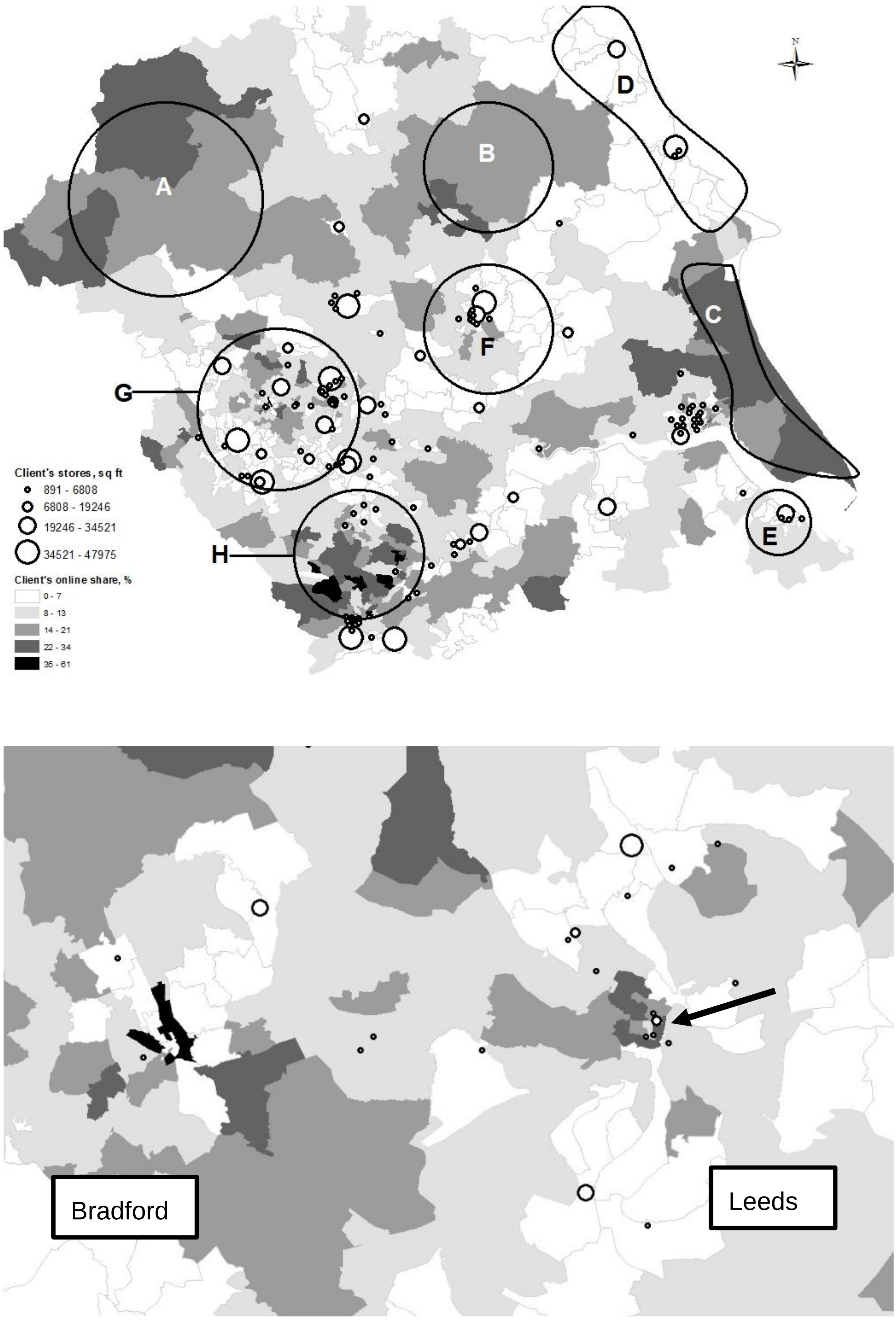


Figure 8.3. Client's Grocery Stores in comparison to online share

Figure 8.3 focuses on client's e-commerce sales plotted alongside the stores of client's only with the highest concentration of stores naturally around the large cities of Leeds, Bradford and York (areas G and F). Client's stores are largely absent in rural areas A and B and these areas have the highest on-line expenditure of up to 34% per postal sector, which demonstrate the substitution between physical and virtual channels and substantiates the efficiency hypothesis which was discussed earlier in the thesis. Coastal areas D and E also have low on-line sales and access to an extensive network of client's stores. The high market share for e-commerce for client in Area C, however, seems to be more difficult to explain. Here, client's on-line share is high despite the presence of a number of large client's stores. Sheffield (area H) is also interesting. Here there are some postal sectors with a very high on-line expenditure (up to 61% in some areas). These tend to be the more affluent western suburbs but for client the majority of their stores in Sheffield tend to be smaller convenience stores. Hence maybe the higher Internet sales are a substitution for a lack of access to the client's major supermarkets.

The enlarged map of the Leeds and Bradford area also shows some interesting patterns. In the northern suburban areas of Leeds, close to a major superstore of over 30,000 sq ft there is a low on-line market share. In contrast, the highest on-line shares are in certain southerly urban areas of Bradford and west/north-west Leeds where there is a limited presence of client's stores. The city centres of both cities also show high levels of e-commerce usage. Consumers in these areas are more likely to belong to the City Living OAC category which was discussed earlier in the thesis as having high overall e-commerce usage. Note too how much of East Leeds and inner north Bradford have low e-commerce usage even though there are plenty of (small) client's stores. Generally these areas do not have access to many superstores – many could be labelled food deserts (Wrigley 2002, Clarke et al 2002). However, geodemographic may be the overarching explanatory factor – many consumers falling into the categories of constrained by circumstance or multicultural. This reminds that the demographic profile of on-line users is important in e-grocery purchasing decision making and accessibility is not the only major factor which has an impact on e-commerce activity.

The analysis to date has identified four major issues in the relationship between on-line share, geodemographics and store provision:

1. High on-line share and high store provision due to demographic profiles of on-line customer (social class ABC1) and possible brand loyalty preferences
2. High on-line share and low store provision due to restricted accessibility to food stores
3. Low on-line share and high store provision. This scenario is interchangeable with the second factor and relates to grocery store accessibility as a major factor in the preference towards on-line spending.
4. Low on-line share and low store provision. This situation is expected in two instances. First, in the "food desert" areas with low grocery store accessibility and a less affluent population. Secondly, in the areas where customers have different brand preferences and use competitors websites to purchase online groceries.

This section has presented one of the first major analyses of actual e-commerce sales for a major UK grocery retailer. This data shows some interesting spatial patterns. On the one hand, there is clear evidence that geodemographics and urban density are important, as found in many other survey based analysis of e-commerce consumption activity. Geodemographic analysis of e-grocery shoppers found greater evidence in support that primary on-line grocery shoppers come from higher social class backgrounds – and are more likely to be rural than urban (in percentage terms). However, the quadrant analysis gives rise to other potentially important findings. Strong evidence was found in support of the efficiency theory with the prevalent number of occurrences of on-line spending in areas with lower physical store provision and less urbanisation. There is a clear indication of substitution between on-line and physical channels in areas with limited accessibility to grocery stores. That said, there is also evidence to support the diffusion of innovation theory with young, city dwellers being enthusiastic on-line shoppers despite the greater presence and variety of nearby grocery stores.

Given these findings what are the implications for retailers? In marketing terms perhaps retailers should target more affluent, rural areas more generally when promoting e-commerce. They should also perhaps look at areas where access to their own physical stores is low as there is clear evidence of substitution taking place when access is poor. This relationship between a store network and the company's e-share of the market is fascinating. It poses interesting questions in relation to the impact on e-commerce sales of store opening and closures. It also raises the issue if e-commerce can be integrated to classic store location forecasting models.

8.3. Designing Spatial Interaction Model to include online grocery sales

In the previous section quadrant analysis established the relationship between grocery store provision and on-line sales with clear evidence of a substitution effect between these two channels. Moreover, it was established that geodemographic characteristics of on-line customers have equal importance in on-line demand estimation (as described in Chapter 6). Based on the research findings hitherto, this section attempts to design a model which will incorporate on-line sales as well as face-to-face sales modelling in Chapter 7.

To adapt the existing Spatial Interaction Model (described in Chapter 7) two additional columns (on-line stores) were created – Client's On-line and Competitors' On-line.

To recap, the following indicators were used to check the model results against actual data and national statistical data (i.e. were used to help calibrate the model).

1. *Nectar Expenditure*. Client's total weekly expenditure across different channels derived from loyalty card data (described in Chapter 5).
2. *On-line Expenditure.* On-line weekly expenditure by client's customers. This number is expected to reflect real client's on-line expenditure due to the almost 100% participation in loyalty scheme of on-line customers.

3. *Stores Expenditure.* Total weekly client's expenditure collected at the POS across all the stores.
4. *Nectar On-line Share.* The share of on-line expenditure of total weekly expenditure derived from Nectar card scheme. The value is 7%.
5. *Store's On-line Share.* The share of client's on-line expenditure (£17,652,498) as a percentage of Stores' Weekly Expenditure (£24,415,660), which equals 5%.
6. *On-line Market Share.* The national figure for online channel share in total grocery expenditure with the current value of 4.4% (see Chapter 2).
7. *Client's On-line Market Share.* Client's market share in total on-line grocery expenditure is currently 17%, although, this figure is expected to be lower for the study area due to the lower presence of client's stores in Yorkshire and Humberside (131 out of the total of 1185), in comparison to London and South East areas which are the home regions of the supermarket chain. In terms of floorspace, 6% of the total client's sales floorspace is present in the study area.
8. *Client's Market Share.* Client's market share in total grocery expenditure in the study area based on the total demand (£149,437,845) and Client's Expenditure derived from loyalty card scheme (£17,652,498).

These eight indicators will be used to evaluate the predicted values generated by the SIM. To check the validity of the models statistical correlation was applied between estimated (SIM) and actual (Nectar) data. Building a SIM to incorporate e-commerce is novel and no published work exists attempting to do this. Thus, the next sections proceed very much in the style of numerical experiments often seen in the development of new forms of model in the past (i.e. Clarke M. and Wilson 1983, 1985, in relation to new dynamic SIMs of urban spatial structure). The next sections outline the experiments of the models with introduction of different variables and different use of the model parameters.

8.3.1 Building a SIM for on-line grocery retailing - Trial 1 - 3

This section describes the initial tests of the modified SIM which includes on-line expenditure thus developing a new methodology to predict on-line sales.

The first major issue is the assignment of an attractiveness value for on-line sites. To reflect attractiveness for the on-line channel two virtual grocery floorspace variables were introduced, one for client's and one for the competitors. These values were set in relation to the average physical store size - 6300sq ft. Then this value was divided between the two on-line 'stores' (the client and the competition) based on their on-line grocery market shares - 17% client's and 83% competitors' (Figure 2.13). The attractiveness values were thus set as 1030 sq ft for client and 5230sq ft. for the combined competition. The initial distance from each postal sector to both on-line stores in the distance matrix was set as 1.45km. This value was reached by significant numerical experiment. This value of 1.45km gave a realistic allocation of expenditure to the on-line stores: the national on-line market share of 4.4%. Table 8.4 shows the SIM results in comparison to actual data (nectar data and store sales data)

Table 8.4. SIM results – Trial 1

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,846,376	6.8
On-line Expenditure (£)	1,214,067	1,144,357	-5.7
Stores Expenditure (£)	24,415,660	18,846,376	-22.8
Nectar On-line Share (%)	7.0	6.1	
Stores On-line Share (%)	5.0	4.7	

On-line Market Share (%)	4.4	4.5	
Client's On-line Market Share (%)	17.0	17.0	
Client's Market Share, %	11.8	12.6	

Negative values indicate an under-estimation in comparison to actual data. For example, actual store expenditure is almost 23% higher in comparison to SIM results. This is largely due to the inability of the faced-to-face model to handle the revenue estimations accurately for convenience stores (Chapter 7).

This simple first stage model did not produce accurate spatial predictions of e-commerce usage. The second trial tested the effect of distance on on-line expenditure between rural and urban areas. Distance was disaggregated between rural and urban areas based on average population density (2110 persons per sq km)on a sliding scale. The distance was set as follows. Rural areas (density 8.24 to 3641 per sq. km) – 2km, Semi Urban (density 3641 – 7283per sq km) - 0.5km, Urban (density – 7283 – 10925 per sq. km) - 0.1km. The average value of distance for online store is 1.3km for all 791 postal sectors. The SIM results for this model are presented in Table 8.5.

Table 8.5. SIM results – Trial 2

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,477,972	4.7
Online Expenditure (£)	1,214,067	775,953	-36.1
Stores Expenditure (£)	24,415,660	18,477,972	-24.3
Nectar Online Share (%)	6.9	4.2	-2.7

Stores Online Share (%)	5.0	3.2	-1.8
Online Market Share (%)	4.0	3.1	-0.9
Client’s Online Market Share (%)	17.0	17.0	
Client's Market Share, %	11.8	12.4	0.55

This model produces a better spatial fit, with more e-commerce users living in rural areas but now considerably underestimates total on-line expenditure, which indicates that the on-line ‘stores’ are generally not attractive to certain customers in terms of their size and/or travelled distance. To increase attractiveness of the on-line stores, the sizes of the on-line stores were increased to 1700sq ft for client’s store and 8300 sq ft for competitors’ store. These changes to the model are reflected in the results in Table 8.6.

Table 8.6. SIM results – Trial 3.

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,934,841	7.3
Online Expenditure (£)	1214067	1232823	1.5
Stores Expenditure (£)	24415660	18934841	-22.4
Nectar Online Share (%)	6.9	6.5	-0.4
Stores Online Share (%)	5.0	5.0	

Online Market Share (%)	4.0	4.9	0.9
Client's Online Market Share (%)	17.0	17.0	
Client's Market Share, %	11.8	12.7	0.86

Increasing the attractiveness of both on-line 'stores' balanced the on-line expenditure with client's predicted and actual (nectar) on-line expenditure values now converging in the model. At the same time, client's on-line market share stayed the same at 17% which is the national average market share for client's on-line sales.

But what about the geography? Predicting the right amount of overall e-commerce sales is crucial, but it is also important to get the right sort of geography – i.e. that sales to on-line 'stores' are coming from the right sort of locations. The first distribution of predicted on-line expenditure is presented in Figure 8.4

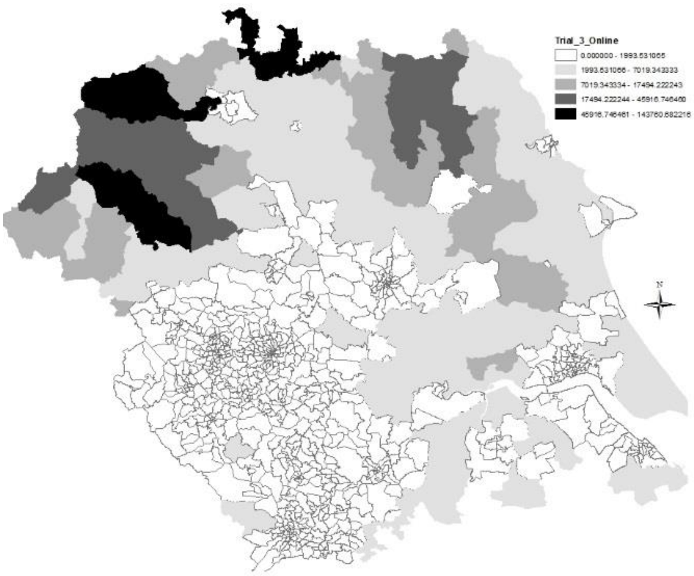


Figure 8.4. Trial 3 –Spatial distribution of online expenditure

Figure 8.4 shows the patterns of on-line expenditure is mostly generated in the northern most rural parts of Yorkshire and Humberside. This is a good start – the evidence presented in Chapter 6 and again highlighted above, shows how important rural areas are for generating on-line sales. However, the model does not capture on-line expenditure in urban areas very well. This problem will be addressed in the next sections with introduction of new variables into the SIM.

8.3.2. Spatial Interaction Model – distance parameter

This section will explore the effect of distance on on-line expenditure in more detail. Distance is a proxy for accessibility as described in Chapter 7. In the first trials the distance to e-commerce stores was set at 0.1km for urban areas, 0.5km for suburban areas and 2kms for rural areas. What this means in reality is that each urban postal sector has a virtual store 0.1km away, but this is a very small virtual store. Hence it is not very attractive to such urban residents as there are many physical stores nearby. For rural areas the opposite is the case. Each rural postal sector has a virtual store only 2km away and although these are again small in size they can be the closest store in that postal sector. Hence the model unsurprisingly allocates a lot of rural residents to e-commerce and few in the more tightly packed urban areas.

To get more realistic distance decay factors in the model it is useful to do some further analysis on the data. The results of analysing the key features of the distance matrix show that average minimum distance to any physical store is 1.4km which is likely to be a competitor's store given their sheer number. The average minimum distance to client's store is 5km with the larger store formats being even further at 7.4km. The closest distance to a client's convenience store is 0.01km. Table 8.7 below summarises the key distances involved.

Table 8.7. Applied distances in SIM

Average minimum distance to store	1.4
Average minimum distance to client's store	5.0
Average minimum distance to client's store over 3500sq ft	7.4
Average minimum distance to client's store under 3500sq ft	0.01
Average maximum distance to client's store	49.0
Average maximum distance to Competitors' store	14.9
Average distance to Competitors' store over 3500sq ft	1.41

A further multiple regression statistical analysis of on-line expenditure and various attributes identified significant variables which include distance to client's stores and distance to the stores of over 3500sq ft. (Table 8.8).

Table 8.8. Multiple Regression – Actual Online Expenditure

Model	Unstandardised Coefficients		Standardised	t	Sig.
	B	Std. Error	Beta		
(Constant)	4.885	5.626		0.868	0.385
Distance to client's stores	0.717	0.074	0.594	9.695	0
3500-10000 sq ft	-0.256	0.085	-0.156	-3.006	0.003
over 30000 sq ft	-0.315	0.073	-0.259	-4.299	0
% AB	0.149	0.039	0.201	3.804	0
% E	0.735	0.182	0.361	4.037	0
Unemployed	-0.224	0.13	-0.139	-1.729	0.084
Full time student econ active	-0.395	0.204	-0.106	-1.938	0.053
Retired	-0.25	0.099	-0.129	-2.513	0.12
Urbanisation	-296.91	159.743	-0.088	-1.859	0.063
Client's provision	-0.241	0.077	-0.117	-3.13	0.002
Competition provision	0.115	0.034	0.129	3.321	0.001

The results of the multiple regression shows that on-line usage decreases with the close proximity of the physical store. Moreover, there is confirmation of the negative relationship between population density and on-line expenditure, i.e. more urbanised areas have more choice of grocery stores in the vicinity with physical stores being a more preferential choice for grocery shopping.

Based on these factors and population density the distance to the nearest on-line virtual store was re-calibrated to be 0.01km (average minimum distance to the nearest physical store)in the urban areas and 14.9km (maximum average distance to the nearest physical store) in the most rural area. In 105 postal sectors with density (8.24 – 100) distance is 10km. In 220 postal sectors with density (100-1000), distance is 5km. In 238 postal sectors with density (1000-3000), distance is 0.5km. In 229 postal sectors with population density of over 3000 people per km the distance to the nearest online store is 0.01km. The results of these changes to the model are presented in Table 8.9.

Table 8.9. SIM results – changes to distance

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,493,683	4.8
Online Expenditure (£)	1,214,067	791,664	-34.8
Stores Expenditure (£)	24,415,660	18,493,683	-24.3
Nectar Online Share (%)	6.9	4.3	-2.6
Stores Online Share (%)	5.0	3.2	-1.7
Online Market Share (%)	4.0	3.1	-0.9

Client's Online Market Share (%)		17.0	
Client's Market Share, %	11.8	12.4	0.6

The model underestimates on-line expenditure as the average distance to the nearest on-line virtual store has increased from 1.4km to 3.5km, although, the client's on-line market share remains about right at 17%. The spatial distribution of client's predicted on-line market share across the study area has however improved, as shown in Figure 8.5. The equal breaks have been changed to match the client's actual online expenditure (see Figure 8.2).

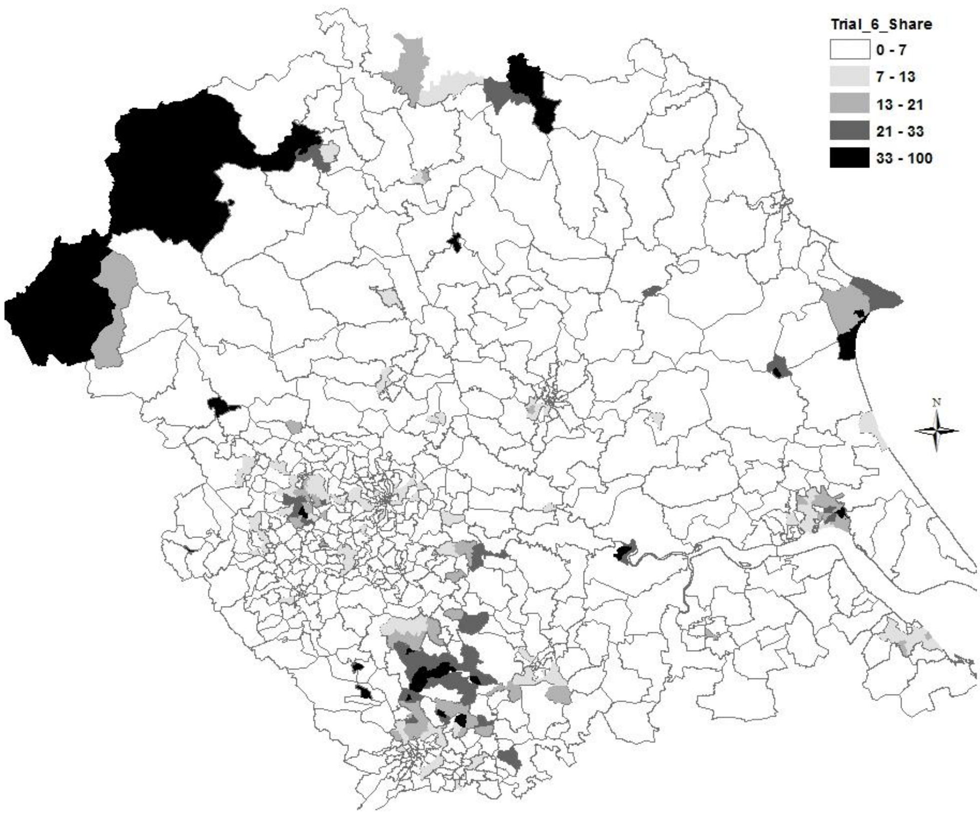


Figure 8.5. Distribution of predicted on-line market share with model version six

The model has started to improve the spatial variations in the estimations across the study area with a better split now between rural and urban areas. However, there is still no account of variations in geodemographics – a key factor identified in Chapter 6 and the analysis above as determining greater e-commerce usage. The next step is to disaggregate the model by demographic characteristics of online and face-to-face customers.

8.3.3. Spatial Interaction Model for on-line and face-to-face: adding demographic parameters

The importance of the socio-demographic characteristics of on-line customers has been outlined in Chapter 3 and 6. The question now is how to quantify the preferences towards on-line channels among various consumer types. In Chapter 4 the role of the alpha parameter in SIM was described, and was further explored in Chapter 7. This parameter can be used to modify the attractiveness of the store (floorspace) to reflect relative attractiveness of one store in terms of its brand, fascia and type over another, by consumer category. In this research consumer preference towards the stores is disaggregated based on their membership of one of the ACORN categories. To begin with the alpha values for the two virtual on-line stores were set to be the same as for the physical stores described in Chapter 7. Table 8.10 presents alpha values chosen for the two virtual stores

Table 8.10. Alpha values for two online stores

Acorn	Online Client's	Online Competitors
1	1.2	1.1
2	1.1	1.05
3	1	0.95
4	0.8	0.85
5	0.7	0.75

Thus, for ACORN group 1 (highest income groups) the on-line virtual stores were set to be more attractive compared to the values for residents in ACORN group 5. This was designed to make on-line stores more attractive for higher income consumers no matter where they reside in the study area. Table 8.11 shows the model results using these additional alpha values.

Table 8.11. SIM results – disaggregation by alpha parameter

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,585,931	5.3
Online Expenditure (£)	1,214,067	920,263	-24.2
Stores Expenditure (£)	24,415,660	18,585,931	-23.9
Nectar Online Share (%)	6.9	4.8	-2.1
Stores Online Share (%)	5.0	3.6	-1.4
Online Market Share (%)	4.0	3.6	-0.4
Client's Online Market Share (%)		16.6	
Client's Market Share, %	11.8	12.4	0.6

Table 8.11 shows the results of the model disaggregation by alpha values. These changes increased estimated on-line expenditure and on-line market share to the national value of 4.4% and reduced underestimation of online expenditure by 10%.

To improve the model further a new parameter K_i was introduced to reflect the combination of multiple attributes of typical on-line grocery shopping demographic characteristics.

$$S_{ij}^{m} = O_i^m A_i^m W_j^{\alpha^m} K_i \exp(-\beta^m d_{ij}) \quad (8.1)$$

where K_i is an additional attractiveness term to boost attractiveness towards online channels in residence zone *I* based on the combination of six demographic characteristics:

1. Family with no dependent children
2. Family with two dependent children with youngest 5-18
3. Social class AB
4. Social class C
5. White ethnicity
6. People aged 45 to 54

These variables were determined by combining two new demographic analyses. First, based on analysis in Chapter 3, multiple linear regression identified the significant demographic variables in relationship to online expenditure (CACI data) with 20 out of 47 variables (see Table 3.2). Secondly, using the Minitab statistical application these six variables were identified as the variables which have an effect on on-line expenditure parameter, the technique designed by Hood (2016). By applying the scoring system each postal sector received scores 1 to 6 based on the number of variables which received a score 1 or 0 (above or below the average value of this variable across the study area). Table 8.12 provides the sample of the combined score analysis data.

Table 8.12. Sample of score analysis

Postal Sector	Family 0 dependent kid	Family 2+ dependent kids youngest 5-18	AB	C1	White	Age 45-64	Score
BD4 7	1	1	1	1	1	1	6
BD4 8	1	1	1	1	1	1	6
BD4 9	1	1	1	1	1	1	6

BD5 0	1	0	0	1	1	1	4
BD5 7	1	1	0	1	1	1	5
BD5 8	1	1	1	1	1	1	6

Based on the scoring system each postal sector *i* received attractiveness parameter represented in the Table 8.13

Table 8.13. Attractiveness parameters based on score analysis

Score	K_i
6	1.1
5	1.05
4	1.02
3	1.01
2	1.005
1	1.005
0	1

Results of the introduction of new K parameter into the model is presented in the Table 8.14

Table 8.14. SIM results – disaggregation by K parameter

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	18,899,522	7.1
Online Expenditure (£)	1,214,067	1,197,503	1.4
Stores Expenditure (£)	24,415,660	18,899,522	-22.6

Nectar Online Share (%)	6.9	6.3	-0.5
Stores Online Share (%)	5.0	4.9	-0.1
Online Market Share (%)	4.0	4.8	0.8
Client's Online Market Share (%)		16.8	
Client's Market Share, %	11.8	12.6	0.8

The new parameter has improved the model with a very good match now between e-commerce predicted use and actual client's online revenue across the study area. Figure 8.6 shows the distribution of the predicted online share with introduction of new parameter K.

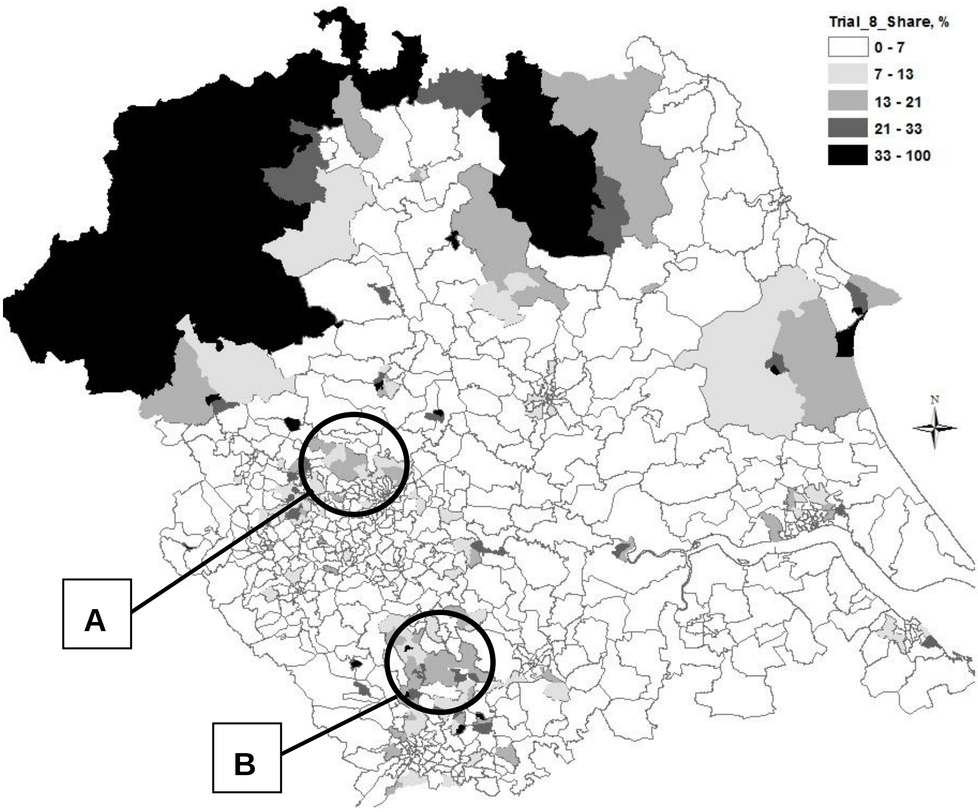


Figure 8.6. SIM results - Disaggregation by parameter K

The model has started to show the spatial distribution of online expenditure between the rural and urban locations. As expected the higher online expenditure is generated in rural areas with less accessibility to physical stores and suburban more affluent areas of Leeds and Sheffield (areas A and B).

8.3.4. Disaggregation by beta value

The final step is to improve the fit between rural and urban postal sectors. Although the right sort of patterns are being generated by the model there is still too much on-line activity predicted in rural areas, despite the previous beta disaggregations. Traditionally β is the parameter which allows customers travels further to the store based on customers' locations. In section 8.3.2 distance in distance matrices to on-line stores was allocated based on the area's population density (urban via rural locations). Consequently, the model has been disaggregated by accessibility with customers living in rural areas having longer distances (up to 10km) to travel to the on-line virtual store, whereas, the city dwellers, effectively, have the on-line virtual stores on their door step. The literature review in Chapter 2 revealed the effect of physical stores and their proximity on on-line expenditure and the quadrant analysis (section 8.2) confirmed the substitution effect between the two channels. The question is how to quantify the effect of physical stores and their proximity to on-line expenditure? The suggestion here would be that customers living in postal sectors with a large client's stores on their door step will be less attracted to the online channel. However, client's customers living in postal sectors with close proximity to only small convenience stores will be more attracted to the on-line channel.

Initially, similar beta values as in the face-to-face SIM (Chapter 7) were applied based on the proximity and size of the store. Now the beta will be varied in relation to access to client's stores (although we do this via the beta parameter we could add a separate accessibility score here). The lowest beta value of 0.2 for on-line was assigned to 88 postal sectors where there is a presence of client's small stores having a floorspace from 1000 to 6000sq ft; 0.4 beta value was assigned to 8 postal sectors with the presence of client's supermarkets – 13000 to 16000sq ft.; 0.5 beta value received postal sectors (11 in total) with client's large stores ranging from 23000sq ft. to 55000sq ft. The postal sectors with no presence of client's stores received a beta value of 1.

In this trial distances have been changed to reflect finer variance between rural and urban areas with following parameters. The most rural areas (density 1 to 100 inhabitants per sq km) received value of 15km. Postal sectors with population density between 100 and 500 per sq km received distance of 10km. The semi-rural areas with population density of 500 to 1000 were assigned 0.5km distance. The urban areas with population density between 1000 and 3000 inhabitants per sq km received 0.1km. The most urban postal sectors with population of over 3000 inhabitants per sq km received 0.01km distance value.

The results of the model disaggregated by beta values and new distances are presented in Table 8.15

Table 8.15. SIM results – disaggregation by beta value

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	19,126,008	8.3
Online Expenditure (£)	1,214,067	1,248,383	2.8
Stores Expenditure (£)	24,415,660	19,126,008	21.7
Nectar Online Share (%)	6.9	6.5	-0.4
Stores Online Share (%)	5.0	5.1	-0.1
Online Market Share (%)	4.0	4.9	0.9

Client’s Online Market Share (%)		17.1	
Client’s Market Share, %	11.8	12.8	1.0

The model predicts an accurate on-line level of usage across the study area with client’s market share matching the national figures of 17%. The total on-line channel market share is 4.9 % which is in the range of the estimated values between 4% and 6% discussed in Chapter 6. The correlation between estimated and actual online expenditure across postal sectors has improved with the value of 51%. The correlation for the total model across all three channels is 83%. Figure 8.7 shows the predicted distribution of on-line market share across the study.

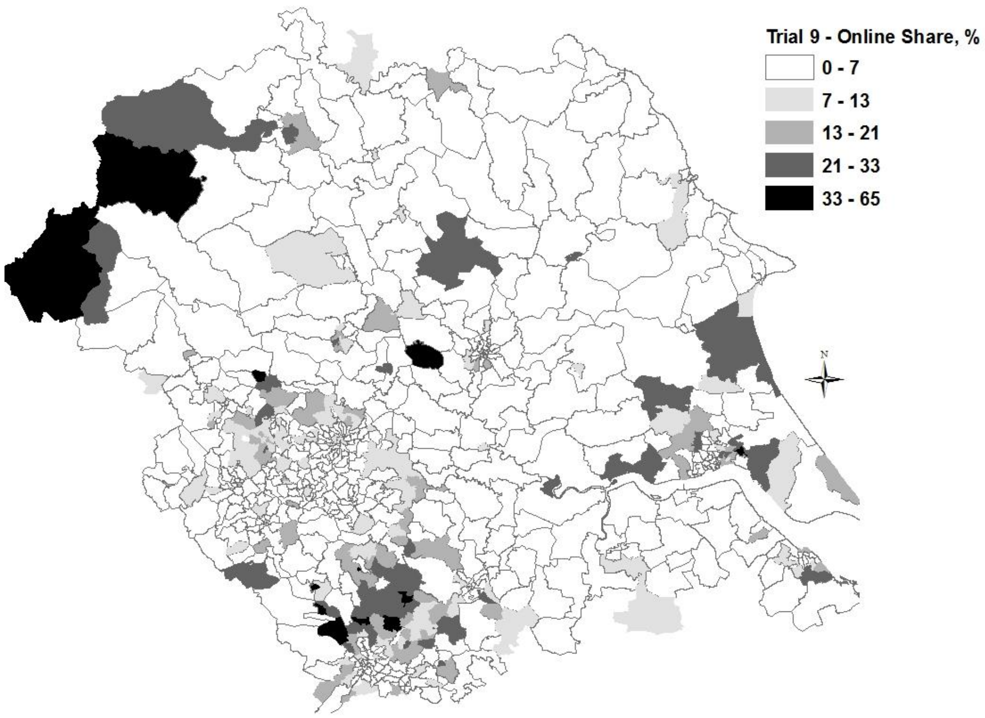


Figure 8.7. SIM results - Disaggregation by beta value

The model shows lower online share in rural areas in comparison to actual data (see Figure 8.2). Although, the spatial distribution of estimated online share and actual values have very similar patterns with the maximum value of on-line share of 65% compared to 61% actual client's online share.

To improve attractiveness of online channel the following changes have been applied. Changed beta value to increase attractiveness towards online channel in the areas with no presence of large client's stores. Beta value of 0.1 received postal sectors with less than 1km to the nearest convenience store. Beta values of 0.7 were applied to the areas with the distance of over 5km to the nearest supermarket of over 3500sq ft. Beta values of 0.9 were applied to postal sectors with the distance to nearest client's supermarket store of over 5km. The results of new changes to the model are presented in Table 8.16.

Table 8.16. SIM results – model calibration

	Actual	SIM	% difference
Nectar Expenditure (£)	17,652,498	19,019,029	7.7
Online Expenditure (£)	1,214,067	1,317,010	8.5
Stores Expenditure (£)	24,415,660	19,019,029	22.1
Nectar Online Share (%)	6.9	6.9	0.0
Stores Online Share (%)	5.0	5.4	0.4
Online Market Share (%)	4.0	4.8	5.1
Client's Online Market Share (%)		16.8	17.3
Client's Market Share, %	11.8	12.7	0.91

Table 8.16 shows the results of model calibration. The model's predicted values stay in line with actual data. Although, the model over estimates client's online expenditure by 8.5%. Figure 8.8 shows the distribution of online share produced by the final model.

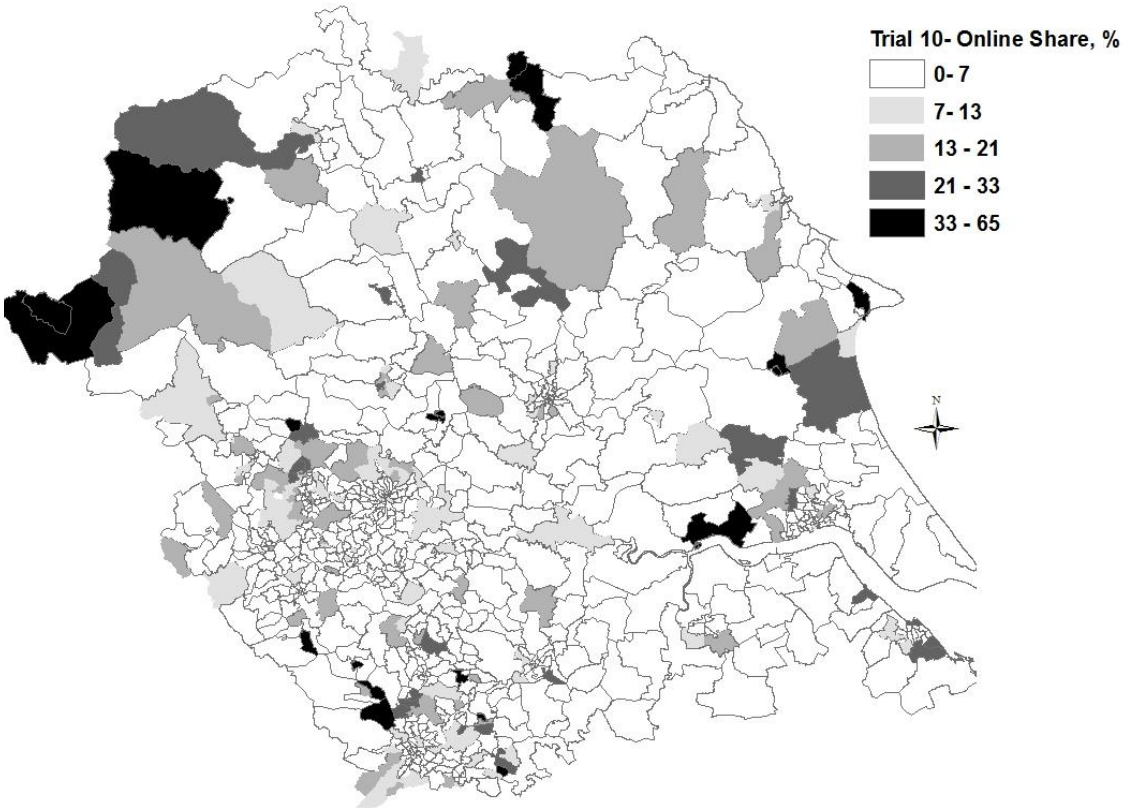


Figure 8.8. SIM results – Final Model

The correlation between estimated and total data has improved by 5% to the value of 55%. The overall model's correlation with actual client's expenditure in the study area is 83%. The correlation between client's stores actual expenditure and estimated data across all postal sectors is 97%. These factors indicate that model has a good level of accuracy of predicting the total values of on-line and in-store expenditure, although, future work is required to improve the model in terms of spatial distribution.

8.3.5 What if scenarios

This section analyses the effect of two scenarios resulting from the proposed development in Yorkshire and Humberside.

1. *Client opens new large store in the area*

The previous section identified the substitution phenomenon between face to face and online channels and this scenario will test its importance under new conditions. The scenario involves opening a large supermarket store of over 20 000 sq ft in the rural area, the north of the study area (DL8 3). Currently, the nearest client's large store of 30 000 sq ft is over 20km away. The only grocery store nearby is a small convenience Spar store with a floorspace of 1740 sq. ft. The area has a high estimated online market share of 45%. The population density is very low with 13 people per sq km. The majority of the households belong to Acorn Category 3 (Comfortable Communities). Figure 8.9 shows the distribution of online share with the new scenario of opening new store in rural area which is marked as 'NS'.

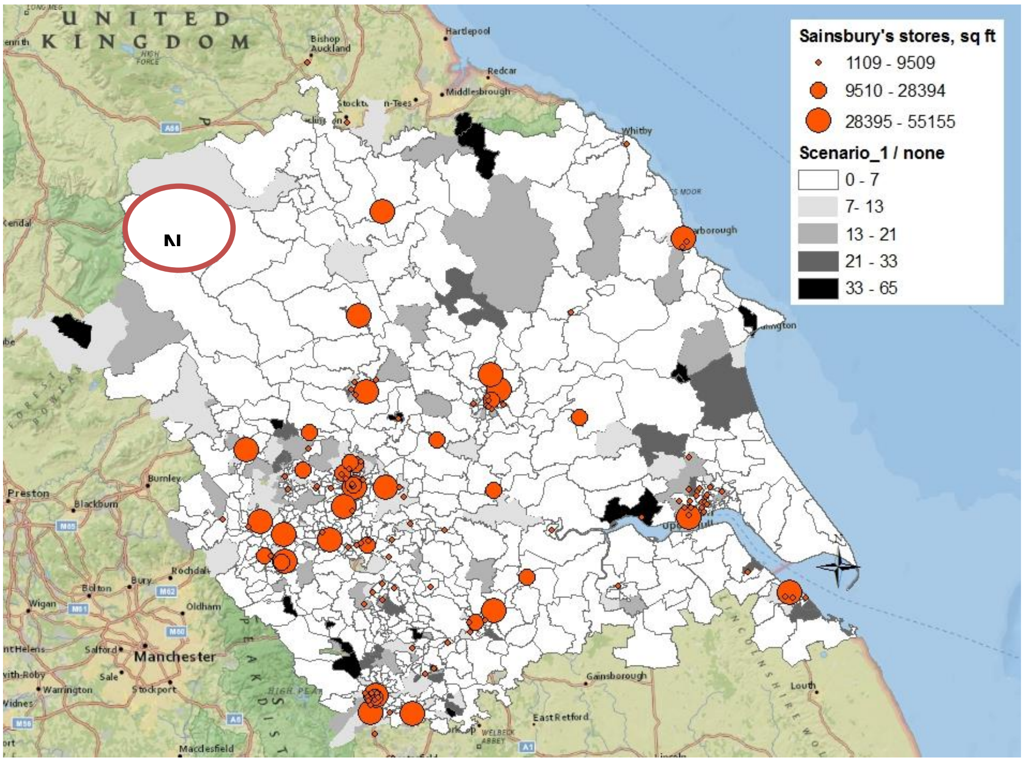


Figure 8.9. Scenario 1 – Opening New Store

The map shows that online market share have changed from up to 45% in the surrounding area to 5% with an opening of new large face to face store. The online channel has become far less attractive with the appearance of a new large physical store in the close proximity. Although this scenario is rather unlikely, it was applied purely to demonstrate how the model can now switch consumers between face to face and online channels. For client they can now appreciate that building a store in a previously open market will have consequences for online sales.

1. *Client closes large face to face retail unit*

In contrast to the first situation, this scenario involves closing the large client's store (almost 40 000sq ft) in a suburban area of Leeds, LS15 9. The nearest large client store is approximately 3km away. Figure 8.10 shows the distribution of online share before the scenario and after. The arrow points to the store which was closed.

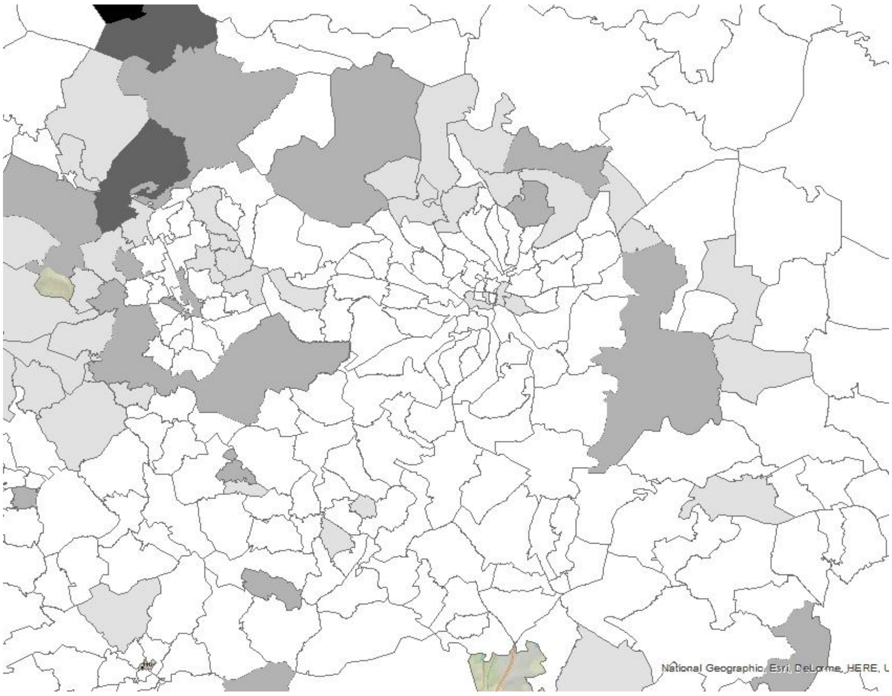


Figure 8.10. Scenario 2 – Closing Client's Store

The map shows that the closure of a large physical store did make online channel more attractive but not by very much. In this instance, the closure of this store made other large physical stores more attractive with share of competitors physical stores increasing by 5%. This shows that the switching effect between face to face and online channels will be less in semi-urban and urban areas as much of the revenue loss would be picked up by face to face stores in the immediate vicinity.

8.3.6. Future improvements

The previous section established that distance to 'online' store, attractiveness towards online channel among various demographic groups and the proximity of physical stores have an effect on online expenditure. To refine the model further the next step would be to disaggregate the model by types of online consumers based on the quadrant analysis and apply consumer clustering technique.

The following types were identified based on the average values of Client's Online Share (OS), Client's Store Provision (SP), Competitors Store Provision (CP), Urbanisation (UR). High values are the values which fall above average and low values are the values which fall below the average values. All 12 consumer types can be aggregated into three major categories as follows:

Client's store provision. These Client's online customers prefer online channel despite having physical stores in close proximity and/or living in urbanised areas

1. CLIENTURBAN–High OS, High SP, High UR
2. CLIENTRURAL - High OS, High SP, Low UR
3. CLIENTPUREURBAN - High OS, Low SP, High UR
4. CLIENTPURERURAL - High OS, Low SP, Low UR

Client's Loyal Online Users. These online customers remain loyal to client's brand despite good grocery provision nearby offered by competitors.

5. PUREURBAN – High OS, High SP, High CP, High UR
6. PURERURAL - High OS, High SP, High CP, Low UR
7. LOYALONLINEURBAN - High OS, Low SP, High CP, High UR
8. LOYALONLINERURAL - High OS, Low SP, High CP, Low UR
9. SUBSTITUION URBAN – High OS, Low SP, Low CP, High UR
10. SUBSTITUION RURAL- High OS, Low SP, Low CP, Low UR

Competition Online Users. These online consumers prefer to buy groceries from the competitors.

11. COMPURBAN -Low OS, Low SP, Low CP, High UR
12. COMPRURAL - Low OS, Low SP, Low CP, Low UR

Moreover, the model needs to explore in more detail the effect of the competitor's online 'stores' on online sales. Currently, the model includes only one online competitor 'store' and other retailers offering online grocery service need to be included more fully to reflect customer's preferences. The beta value for online competitor online 'store' needs to be refined further to reflect attractiveness of competitor's physical stores including brand preferences and not only the size which will offer more accurate distribution of online expenditure between client's and competitors online 'stores'.

8.4. Conclusions

The aim of this Chapter was to develop the modelling technique which includes online expenditure based on the SIM developed in Chapter 7. The developed SIM tested the effect of three factors – demographic characteristics of online customers, proximity and size of physical stores and customers locations (rural via urban). The various trials of the model established that these factors have an effect on online expenditure. For example, rural locations generate more online expenditure and areas with close proximity of large supermarket stores generate less online expenditure. The model prediction level is relatively good at 55% correlation between actual and estimated online expenditure in the study area. Given a large small number problem here that is a very promising level of correlation and the spatial patterns now look convincing. Moreover, the model correctly produces client's online market share and overall market share of online channel in the study area. Although, the spatial distribution of online share across the postal sectors still shows underestimation in some areas and in other areas the model over predicts online expenditure. There is always a random element to shopping patterns and another set of data to calibrate on might show that some of these poorly performing postal sectors are anomalies of a certain time period. However, to improve the spatial distribution of online expenditure further improvements will be required, First, more detailed demographic classification of online customers needs to developed. Secondly, the model needs to include more detail on other competitors offering online grocery service to reflect the preferences among various demographic groups towards online channel. Further thoughts on model improvements will be given in the concluding chapter.

Chapter 9: Discussions, conclusions and future research development

9.1. Introduction

The principal aim of this thesis was to build a model to estimate grocery expenditure including online sales for use in retail location planning. The Spatial Interaction Model (SIM) is a commonly used technique in location planning for estimating grocery expenditure and has a history of producing estimates of revenues with great accuracy. However, it is important to continually monitor its suitability to handle changing, often more complex consumer behaviour. The recent works of Newing (2013) and Hood (2016) have demonstrated that the classic SIM can be modified to include seasonal visitor demand and non-residential demand which is increasingly important for estimating revenues at certain central city convenience stores. The research presented in this thesis has effectively fulfilled the aims outlined in Chapter 1: to develop new spatial modelling techniques that allow retailers to first understand the geography of online sales and then to include online sales in models of store forecasting. The outcomes of this research will enhance the location planning process in the grocery retail sector, especially of the client's grocery stores.

There are two major outputs of this research. First is the construction of a new demand layer for online sales at the small area geography level (postal sector level). Secondly, the research has developed a model which can be used to predict consumer flows and estimate retailers revenues and market shares in both channels – face to face and online. The outcomes of this research were based on the characteristics of the study area – Yorkshire and Humberside and produced at the postal sector level of geography. There is a nothing to prevent this methodology being replicated in any other location to estimate grocery online expenditure. The model produced in this thesis has also been able to predict physical store revenue also with great accuracy. The model was disaggregated on the demand and supply sides offering a more advanced approach for estimating grocery demand in both channels (online and offline) and it is a powerful tool for modelling spatial patterns of consumer flows, store revenue and market shares. This research was made possible as a result of a close collaboration with leading supermarket chain and CACI, the partner organisations. The model results were calibrated against actual sales data provided by client's planning team. The following sections outline the summary of research findings

9.2. Summary and analysis of research findings

This section provides a discussion of the research outcomes in relation to the proposed objectives outlined in Chapter 1. The aims of this research were developed with regards to existing difficulties in the location planning process for new store development and grocery estimation demand, in relation to online activity. This section also provides some limitations of the data sources and research findings.

9.2.1. Understanding of geography of e-commerce activity

To achieve the main aim of this research (to develop a model to include online sales) it was necessary to identify the major drivers of online sales. First, Chapter 2 provided an extensive review of existing information and theories with regards to the nature of online shopping. Moreover, Chapter 2 introduced the UK e-commerce industry and outlined its future developments with particular attention to the UK grocery industry, its structure and the problems it encounters. E-commerce activity for the purpose of this research was defined as a commercial activity performed on the Internet between businesses and consumers. The important phenomenon of omni-channel retailing was closely examined with a discussion of the changing shopping process and changing consumer behaviour. Furthermore, the effect of face to face channel on online and vice versa was analysed through a detailed literature review identifying three major effects – complementarity, substitution, modification and neutrality.

Next, various studies on online consumer segmentation were examined (with closer attention paid to grocery shopping) in order to understand the characteristics of the typical online consumer. Many scholars have argued that gender, age and social class are the major characteristics of online customers. To identify the geography of online customers, two theories were examined – the diffusion of innovation and the efficiency theory. Both of these theories were tested using the actual online sales data provided by leading supermarket chain.

This demographic segmentation was supplemented by an examination of online customer characteristics present within major geodemographic classifications – in particular the National Readership Survey, CACI's Acorn system and the ONS Output Area Classification. The geodemographic analysis of online customers produced two major outcomes. Whilst, online customers are more likely to be young professional families living in the city or more affluent suburban areas (which supports the diffusion of innovation theory), they also can increasingly belong to the least affluent social categories, i.e. Urban Adversity and Financially Stretched. Interestingly, even poorer pensioners are becoming more enthusiastic online grocery customers which indicates that the online grocery market is maturing and the demographic profile of online shopper is becoming more complex (and less varied in geodemographic terms). The traditional profile of online customer is slowly changing from young professional male (as suggested by diffusion of innovation theory) to the price consciousness consumer who utilises the Internet to maximise their savings. The online customer profile is becoming more analogous to the profile of the general grocery shopper with their own brand preferences. The online channel is becoming an integral part of the overall retailing operation and not simply a satellite unit. Brand preferences will be more important in customers' choice of online retailer as customers becoming more Internet savvy with widespread access to high speed Internet.

The analysis of client's actual data also identified that online customers are still more likely to belong to the higher social classes living in the affluent suburban and rural areas. These findings are not conclusive as client's customers are more likely to belong to AB social classes and reflect the general distribution of the retailers' target market and not necessary all online grocery users. To make conclusive findings on the profile of online customer it would be beneficial to examine actual data of a low cost online grocery retailer, e.g. Asda.

9.2.2. Developing a face to face Spatial Interaction Model (SIM)

The second major aim of this research was to design a SIM based on residential expenditure spent at face to face grocery stores in the study area. SIM has become a key site location modelling technique for retailers, offering a more sophisticated methodology to estimate store revenues in a very competitive, arguably saturated environment. This technique was a natural choice for this research as client's planning team uses this modelling technique to identify locations for their new sites and to estimate store revenues. Moreover, this model reflects consumer behaviour very well and can be modified to include new layers of new consumer attributes (e.g. seasonal variation in demand or, as in the case of this thesis, online sales). The classic production constrained entropy model was applied in this research. The model consists of three major components – demand, supply and interaction. The available residential grocery expenditure in the study area (across 791 postal sectors) was used to construct the demand side of the model. The supply side consisted of the available grocery floorspace of all major grocery stores in the region (1812 in total). The model was further disaggregated to reflect the brand preferences among various Acorn demographic categories. The interaction side of the model relates to store accessibility and travel distance. The parameter beta was refined to reflect the cost of travelling among customers within different locations (rural via urban). The model was calibrated against actual grocery sales data for a period of three months provided by client. The final model produced results close to actual data with a high level of accuracy. The model showed a correlation between actual and estimated grocery revenue data across postal sectors of 83%. Moreover, the model had a close fit between observed and predicted travel distances. The estimated market shares among major grocery retailers were very similar to regional figures based on the available grocery floorspace in the region. It was very important to develop the face to face SIM with a high level of accuracy as the next stage was to add new layers to incorporate online expenditure. It was important that these new layers were added to a well-performing model if the outputs were to be of use to real site location teams.

9.2.3. Developing a new model incorporating online sales

The ultimate aim of this research was to design a model which could predict online sales at the specified level of geography. For online sales, the face to face SIM was modified to include two online 'stores' – Client's and Competitors. The research identified three factors which influenced the level of online expenditure in each postal sector. First, is the effect of the presence in the locality of physical stores. The quadrant analysis undertaken showed the high degree of the substitution effect between the two channels (described in Chapter 8). Second, the demographic characteristics of online customers are also important. The statistical analysis identified the close correlation between online expenditure and the following demographic attributes – social class (AB and C), family composition (families with dependent children), ethnicity (white) and age (45 to 54). Finally, urban/rural is an important component of online sales. To quantify these factors the following modifications were made to the model. A new parameter K was introduced to reflect the preferences towards online channel of customer with various demographic characteristics. The model was disaggregated to reflect preferences towards the two online 'stores' among Acorn categories. To quantify distance to virtual stores the optimal values were identified based on location type (rural, semi-rural, suburban and urban) with the furthest distance set at 15 km for customers living in a very remote location with less than 100 people per sq km. Postal sectors closer to the urban central areas had a very low distance to online 'stores'. The rationale for this reflects the fact that in very rural areas a low travel distance to online 'stores' would make them too attractive in the model, given the fact that there are few face to face stores in the vicinity. Conversely, in built up urban areas the accessibility to face to face stores is high meaning distance travelled to online 'stores' needs to be low to attract customers to them. The exact values were determined by calibrating online sales to match client's data.

To further model the proximity and attractiveness of physical stores in comparison to online ‘stores’ (size) the model was disaggregated by beta value (the balancing parameter for travelled distance), building on the face to face model. Again the beta values were calibrated in line with actual client’s data. The multiple trials were completed with various parameters (a set of numerical experiments). The final model showed a correlation between actual and predicted online expenditure of 51%. Although that seems low, the small number effect is important to consider – predictions of 2 consumers using online channels when only 1 actual person is in the data can produce seemingly poor fits (100% difference here).

The model’s results were thus tested against various parameters – client’s market share, national grocery online channel market share and client’s actual online expenditure. The model produced a good overall fit in terms of client’s actual online share. The model predicted 16.8%, very close to the actual figure of 17%. Similarly, the combination of parameter values produced a final market share for the online channel at 4.8% compared to 4.4%, the national figure. The spatial distributions of predicted online sales showed the variance between rural and urban areas as expected. Overall, the model at this stage can be considered as reasonably successful given it has been the first attempt to quantify and predict online sales in a SIM. Given the small number problems in many postal sectors, the final geography of the online estimations does capture actual sales well – especially given the three main driving factors described above.

9.3. Further development and future research

The research reported in this thesis attempted to establish the relationship between sales at physical and online stores. Moreover, the research introduced new variables into a SIM to reflect major attributes of online grocery shopping. These parameters chosen help to make a start in producing a new workable SIM which includes online activity. Clearly, further research is required to develop alternative parameters and perhaps continue to refine the variables chosen here. As with all modelling techniques, similar results could be obtained by refining existing variables and parameters further or by introducing entirely new terms in future models, e.g. car ownership. The distance terms and demographic variables used here are not exhaustive. It is hoped other researchers can build on the work presented here to produce even better fitting models in the future.

The modelling itself can be enhanced perhaps with introduction of new techniques. For example, an interesting future project would be to fuse SIM with agent-based models to create a set of individual SIMs. Another consideration for a future project would be to include 'click and collect', as it is different in nature to pure online orders and creates different shopping environments, and hence interesting new geographies to explore.

The model needs to be updated with changing consumer attitudes and social trends. As it was noted previously, online retailing is becoming an integral part of retail operations and is becoming more integrated into an omni-channel retail channel mix, seen less today as a free standing operation. Consumers expect firms to run smooth online operations and offer greater online services. Online channels will undoubtedly become more important and model parameters need to be constantly revisited in order to reflect that.

It is hoped that the new SIM presented here, including both a face to face and online component, will offer new insights for client's store location team. Understanding the geography of demand for online retailing is powerful in its own right. The ability to model the impacts of store openings and closures (both client's and the competition) on not only face to face but now online sales will offer a very powerful tool for revenue estimation in the future.

List of References

Allemand, S. 2001. Les enjeux des mobilités quotidiennes. *Sciences Humaines*, June, pp.46-62.

Alexander, A., Cryer, D., & Wood, S. 2008. Location planning in charity retailing. *International Journal of Retail & Distribution Management*, 36(7), pp.536-550.

Alvarez and Marcel. 2014. *Home delivery fulfilment in UK Grocery.* [Accessed May 4th 2015]. Available from: http://www.alvarezandmarsal.co.uk/sites/default/files/sidebar-callouts/am_ecs_homedeliveryfulfilment_f2_web_2.pdf

Anderson, M., Hagen H., Harter G. The Coming Wave of "Social Apponomics" [Accessed April 4th 2013]. Available from: https://www.strategy-business.com/article/11101?gko=6092d

Batty, M. and Macky S. 1972. The Calibration of Gravity, Entropy and Related Models of Spatial Interaction. *Environment and Planning*. 4, pp.377-405

Bhatnagar, A. and Ghose, S. 2004. Segmenting consumers based on the benefits and risks of Internet shopping. *Journal of Business Research*, 57(12), pp.1352-1360.

BBC News. 2013. *High Street chain store closures soar, says research*. [Accessed March 5th 2016]. Available from: http://www.bbc.co.uk/news/business-21611772

Birkin, M., Clarke G., Clarke, M. 2002. Retail geography and Intelligent Network Planning. Chichester. John Willey & Sons.

Birkin, M., Clarke, G. And Clarke, M. 2010a. Refining and operationalising entropy maximising models for business applications. Geographical Analysis, 42(4), pp.422-445.

Birkin, M., Clarke, G., Clarke, M. And Wilson, A. 2010b. The achievements and future potential of applied quantitative geography: a case study. Unpublished working paper: School of Geography, University of Leeds. Copy available from: martin.c.clarke@btInternet.com.

Bradshaw T. 2011. Funds to help bridge UK digital divide. [Accessed May 7th 2015]. Available from: https://www.ft.com/content/2287b23c-c7ea-11e0-9501-00144feabdc0

Bonabeau, E., 2013, December. Big Data and the bright future of simulation—The case of agent-based modeling. In *Simulation Conference (WSC), 2013 Winter* (pp. 1-1). IEEE.

Burt, S., Sparks, L. and Teller, C. 2010. Retailing in the United Kingdom – a Synopsis. *European Retail Research*, 21 (1), pp.173-174.

CACI. 2014. *ACORN User Guide.* [Accessed May 7th 2015]. Available from: http://acorn.caci.co.uk/downloads/Acorn-User-guide.pdf

CACI. 2016. *About CACI.* [Accessed March 5th 2016]. Available from: http://www.caci.com/about.shtml

CACI. 2016. *Online Grocery Shopping is Evolving.* [Accessed July 4th 2016]. Available from: https://www.caci.co.uk/sites/default/files/imce/Online_grocery_stores.pdf

CACI. 2013. *CACI Spend Estimates and Projections User Guide.* Centre for Retail Research. 2016. *Online Retailing: Britain, Europe, US and Canada 2016.*

Clarke, I., Mackaness, W. and Ball, B. 2003. Modelling intuition in retail site assessment (MIRSA): making sense of retail location using retailers' intuitive judgements as a support for decision-making. *International Review of Retail, Distribution & Consumer Research*, 13(2), p.175.

Clarke, I., Kirkup, M. and Oppewal, H., 2012. Consumer satisfaction with local retail diversity in the UK: effects of supermarket access, brand variety, and social deprivation. *Environment and Planning A*, *44*(8), pp.1896-1911.

Clarke G. *et al*, 2015. [Accessed June 10th 2016]. Available from: http://www.retailresearch.org/onlineretailing.php

Clarke G., Birkin . 2016. Forthcoming

Cliquet, G. 1990. La mise en oeuvre du modèle interactif de concurrence spatiale (MCI MODEL) subjectif. *Recherche et Applications en Marketing*, 5, 1, pp.3-18.

Cliquet, G., 2006. *Geomarketing* (No. halshs-00188683).

Competition Commission. 2008. *The supply of groceries in the UK market investigation.* [Accessed April 10th 2014]. Available from: http://www.ias.org.uk/uploads/pdf/Price%20docs/538.pdf

Consultancy UK. 2016. *European B2C e-commerce market breaks through €500 billion mark.* [Accessed June 1st 2016]. Available from: http://www.consultancy.uk/news/12175/european-b2c-e-commerce-market-breaks-through-500-billion-mark

Copeland, M.T., 1923. Relation of consumers' buying habits to marketing methods. *Harvard business review*, *1*(3), pp.282-289.

Couclelis, H., 2004. Pizza over the Internet: e-commerce, the fragmentation of activity and the tyranny of the region. *Entrepreneurship & Regional Development*, *16*(1), pp.41-54.

CRR, 2016. *Online Retailing: Britain, Europe, US 2016.* [Accessed June 6th 2016]. Available from: http://www.retailresearch.org/onlineretailing.php

Datamonitor. 2012. *Online shopping preferences*. [Accessed June 4th 2013]. Available from: https://www.globaldata.com/retail/

Davies, M., and Clarke, I. 1994. A framework for network planning. *International Journal of Retail & Distribution Management*, 22(6), p.6.

DEFRA. 2015. *Food Statistics Pocketbook.* [Accessed June 6th 2016]. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/526395/foodpocketbook-2015update-26may16.pdf

Digital Foodie. 2016. Evolution of Online Groceries. [Accessed June 4th 2016]. Available from: http://www.digitalfoodie.com/evolution-of-online-groceries/

Dion, D., Cliquet, G. 2006. Consumer Spatial Behaviour. In Cliquet G. Ed. *Geomarketing: Methods and strategies in Spatial Marketing* London: ISTE Ltd, pp. 30-57.

Duggal 2007. *Retail Location Analysis: A Case Study of Burger King & McDonald's in Portage & Summit Counties, Ohio.* [Accessed June 6th 2016].Available from:https://etd.ohiolink.edu/rws_etd/document/get/kent1196133312/

Econsultancy 2014. 31% of UK shoppers research in-store before making purchase online: report. [Accessed June 3rd 2015] Available from: https://econsultancy.com/blog/64394-31-of-uk-shoppers-research-in-store-before-making-purchase-online-report/

Economics. 2015. *Broadband statistics in UK and rest of World.* [Accessed October 5th 2015]. Available from: //www.economicshelp.org/blog/6169/economics/broadband-statistics-in-uk-and-rest-of-world/

Econsultancy. 2016. *UK online retail sales to reach £52.25bn in 2015: report.* [Accessed January 27th 2016]. Available from: https://econsultancy.com/blog/66007-uk-online-retail-sales-to-reach-52-25bn-in-2015-report/

Elgar, E. 2008. Trust and New Technologies: Marketing and Management on the Internet and Mobile Media. *In: Teemu K and Heikki K. Ed. Trust and New Technologies: Marketing and Management on the Internet and Mobile Media.* [Accessed August 5th 2014]. Available from: https://books.google.co.uk/books?id=0As92opk0j4C&pg=PA71&lpg=PA71&dq=Risk+profile+and+consumer+shopping+behaviour+in+electronic+and+traditional+channels&source=bl&ots=g2sRbHAyUZ&sig=buFH5H6BeJBeQWpeZFXDegWPzrE&hl=en&sa=X&ved=0ahUKEwjWp6rbnPvNAhUhJ8AKHds9D48Q6AEILzAC#v=onepage&q=Risk%20profile%20and%20consumer%20shopping%20behaviour%20in%20electronic%20and%20traditional%20channels&f=false

EMarketer. 2016. *Worldwide Retail Ecommerce Sales: eMarketer's Updated Estimates and Forecast Through 2019.* [Accessed September 10th 2015]. Available from:http://www.emarketer.com/Report/Worldwide-Retail-Ecommerce-Sales-eMarketers-Updated-Estimates-Forecast-Through-2019/2001716

EMarketer. 2016. *UK Retail Ecoomerce Sales to Reach £60 billion This Year.* Available from: http://www.emarketer.com/Article/UK-Retail-Ecommerce-Sales-Reach-60-Billion-This-Year/1012963

Evolution Insights. 2010. Online Food & Grocery: The shopper prospective. [Accessed September 10th 2013]. Available from: https://www.slideshare.net/evolutioninsights/online-food-grocery-the-shopper-perspective-sample-extract

Eurostat. 2013. *E-Commerce statistics*. [Accessed March 3rd 2014]. Available from: http://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals

Farag, S., 2006. E-shopping in the Netherlands: does geography matter?. *Environment and Planning B: Planning and Design*, *33*(1), pp.59-74.

Fernie, J., Sparks, L., McKinnon, A.C. 2010. Retail Logistics in the UK: past, present and future. *International Journal of Retail & Distribution Management* 2010 38:11/12, pp. 894-914 [Accessed June 5th 2015]. Available from: http://www.emeraldinsight.com/doi/full/10.1108/09590551011085975#_i10

Forengham A.S. and Wong D.W.S. 1991. The Modifiable Areal Unit Problem in Multivariable Statistical Analysis. *Environment and Planning* A. 23, pp. 1025-1044.

Fotheringham, A. S. 2013. Localised Spatial Interaction Models. *In: IGU Conference 2013: Applied GIS and Spatial Modelling, Leeds.*

Goldstein, D. 2007. *What is customer segmentation?* Available from: http://www.mindofmarketing.net/2012/05/what-is-customer-segmentation/

Gupta A., Su B., Walter Z. 2004. Risk profile and consumer shopping behaviour in electronic and traditional channels. *Decision Support Systems*, 38, pp.347-367.

Gong, W., Maddox, L. 2011. Online Buying Decisions in China. *The Journal of American Academy of Business*. 17(1), pp. 43-50.

Gov.UK. 2015. *Broadband delivery UK.* [Accessed March 3rd 2016]. Available from: https://www.gov.uk/guidance/broadband-delivery-uk

Hargittai, E. and Hinnant, A., 2008. Digital inequality differences in young adults' use of the Internet. *Communication Research*, *35*(5), pp.602-621.

Haynes, K. E., Fotheringham, A. S. 1984. *Gravity and Spatial Interaction Models, Sage Publication, Scientific Geography, vol. 2, 1984.*

Heinonen, K. 2011. Consumer activity in social media: managerial approaches to consumers' social media behavior. *Journal of Consumer Behavior*, 10, pp.356-364.

Hernández, T., & Bennison, D. 2000. The art and science of retail location decisions. *International Journal of Retail & Distribution Management,* 28(8), pp.357-367.

Hood, N. 2016. Forthcoming

Humby, C., Hunt, T. and Phillips, T. 2008. *Scoring Points: How Tesco Continues to Win Customer Loyalty.* Kogan Page, London

Hughes R., Hallsworth A. G., and Clarke G. 2009. Testing the effectiveness of the proposed UK 'competition test'. *The Service Industries Journal,* 29(5), pp.569-590.

IGD. 2014. *Store choice vs Product choice: The influence of price.* [Accessed March 8th 2015]. Available from: https://www.igd.com/Research/Shopper-Insight/Channels-and-In-store/25283/Price-what-does-it-mean-for-stores-and-brands-/

IGD. 2015. *Understand your online grocery shoppers.* [Accessed May 3rd 2016]. Available from: http://shoppervista.igd.com/FreeContentView.aspx?cid=16

IGD. 2015. Multichannel trends and innovation in grocery. Available from: http://www.igd.com/Research/Training/Accelerate-your-personal-development/30144/Multichannel-trends-and-innovation-in-grocery/

IGDa. 2016. *The growth potential of online grocery.* [Accessed July 1st 2016]. Available from: http://www.igd.com/Research/Shopper-Insight/The-growth-potential-of-online-grocery/

IGDb. 2016. *Pushing Online Growth Further.* [Accessed April 12th 2016]. Available from: http://www.igd.com/Research/Shopper-Insight/Pushing-online-shopping-growth-further/

Intelligent Positioning. 2013. *Online grocery shopping*. [Accessed June 9th 2015]. Available from: http://www.intelligentpositioning.com/en/

Ispira. 2012. Omni channel retailing. Available from: http://www.ispira.com/upload/c4/c4ca4238a0b923820dcc509a6f75849b/11ccd189d25d9ebea10bd2ff18e618ed.pdf Accessed 5 December 2013

Janssen, M.A., 2005. Agent-based modelling. *Modelling in ecological economics*, pp.155-172.

Javadi, M., Dolatabadi, H., Nourbakhsh, M., Poursaeedi, A. and Asadollahi, A. 2012. An analysis of factors affecting on online shopping behaviour of consumers. *International Journal of Marketing Studies*. 4(5). pp.81-98.

J, Sainsbury plc. 2016. *Nectar is the most popular scheme in the UK.* [Accessed June 9th 2016]. Available from: http://www.j-sainsbury.co.uk/media/latest-stories/2010/20100208-nectar-is-the-most-popular-loyalty-scheme-in-the-uk/

Kantar Worldpanel. 2015. *FMCG online sales to reach $130billion by 2025*. [Accessed 10th September 2015]. Available from: http://www.kantarworldpanel.com/global/News/FMCG-online-sales-to-reach-130-billion-by-2025

Kantar Worldpanel. 2016. Grocery Market Share. http://www.kantarworldpanel.com/en/grocery-market-share/great-britain

Kantar Worldpanel. 2016. *Grocery Market Share (12 weeks ending).* [Accessed March 1st 2016]. Available from: http://www.kantarworldpanel.com/en/grocery-market-hare/great-britain/snapshot/03.01.16/

Konus U., Verhoef P.C., Neslin S.A. 2008. Multichannel shopper segments and their covariates. *Journal of Retailing*, 84(4), pp. 398-413.

Korgaonkar, P. and Wolin, L. 2002. Web Usage, Advertising, and Shopping: RelationshipPatterns. *Internet Research: Electronic Networking and Policy*, 12(2), pp. 191-204.

Lieke van Delft. 2013. *Omni Channel Shopping Behaviour during the customer journey.* Available from: http://nrw.nl/wp-content/uploads/2014/06/Omni-channel-shopping-behavior-during-the-customer-journey.pdf Accessed 15 July 2015

Langston P. 2011. Effect of 'click and collect' service on online channel. Presentation

Longley, P., Ashby, D., Webber, R. and Li, C., 2006. Geodemographic classifications, the digital divide and understanding customer take-up of new technologies. *BT Technology Journal*, *24*(3), pp.67-74.

Longley, P.A., Webber, R. and Li, C., 2008. The UK geography of the e-society: a national classification. *Environment and Planning A*, *40*(2), pp.362-382.

Longley, P. and Singleton A. 2009. *Classification through consultation: public views of the geography of the e-Society.* [Accessed July 24th 2016]. Available from: http://www.tandfonline.com/doi/full/10.1080/13658810701704652?scroll=top&needAccess=true

LS Retail. 2016. *Why smart online retailers are opening physical stores.* [Accessed July 24th 2016]. Available from:http://www.lsretail.com/blog/smart-online-retailers-opening-physical-stores/

Mangold, WG. and Faulds, DJ. 2009. Social Media: the new hybrid element of the promotion mix, *Business Horizons*, 52, pp.357-365.

McClellan, J.2003. Sweet smell of success. *The Guardian*, September, p. 4.

Mercedes Benz. 2015. UK Online Grocery Retailing: Winter Operating Insight. [Accessed April 4th 2016]. Available from: http://tools.mercedes-benz.co.uk/current/vans/pdfs/fleet-business/winter-operation-insight-online-grocery.pdf

Mintel. 2015a. Online Grocery Retailing – UK – March 2015' report. Available from: http://store.mintel.com/online-grocery-retailing-uk-march-2015 Accessed 1st July 2016

Mintel. 2015b. Online Grocery Retailing – UK – March 2015 report. [Accessed August 3rd 2015]. Available from: http://store.mintel.com/online-grocery-retailing-uk-march-2015?cookie_test=true

Mintel. 2016. 29% of UK Online Grocery Shoppers are shopping for groceries more online now than a year ago. [Accessed June 1st 2016]. Available from: http://www.mintel.com/press-centre/retail-press-centre/29-of-uk-online-grocery-shoppers-are-shopping-for-groceries-more-online-now-than-a-year-ago

Mokhtarian, P. 2004. A conceptual analysis of the transportation impacts of B2C e-commerce. *UC Davis.* [Accessed February 3rd 2013]. Available from: http://escholarship.org/uc/item/74m9x4sh

Mokhtarian, P.L. and Salomon, I., 2001. How derived is the demand for travel? Some conceptual and measurement considerations. *Transportation research part A: Policy and practice*, *35*(8), pp.695-719.

Mrs Bargain, Hunter. 2014. Best Grocery supermarket Click and Collect service. [Accessed January 7 2015]. Available from: http://www.mrsbargainhunter.co.uk/shops/grocer-best-click-collect-service

Mutlu, Y. and Tufan, O. 2015. Determining the effects of the Perceived Utilitarian and Hedonic Value on Online Shopping Intentions. *International Journal of Marketing Studies*, 7(6). [Accessed May 7th 2016].

Newing, A., Clarke, G., Clarke, M. 2013. Identifying seasonal variations in store-level visitor grocery demand. International Journal of Retail & Distribution Mnagement. Available from: http://www.emeraldinsight.com/doi/abs/10.1108/09590551311330843 Accessed 6 May 2014

NRS. 2015. Social Grade. [Accessed June 16th 2015]. Available from: http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/

NSR. 2016. Social Grade. [Accessed June 5th 2016]. Available from: http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/

OECD. 2008. Broadband and the Economy. [Accessed May 4th 2014]. Available from: https://www.oecd.org/sti/40781696.pdf

OECD. 2011. OECD Guide to Measuring the Information Society 2011. [Accessed June 1st 2016]. Available from: http://www.oecdbookshop.org/browse.asp?pid=title-detail&lang=en&ds=&ISB=9789264113541

Ocado. 2016. The Evolution of shopping. [Accessed June 23rd 2016]. Available from: http://www.ocadogroup.com/who-we-are/our-story-so-far.aspx

ONS. 2010. British Population Survey. [Accessed June 23rd 2013]. Available from: http://www.thebps.co.uk/

ONS. 2009. National population Projections 2008. [Accessed May 1st 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://ons.gov.uk/ons/rel/npp/national-population-projections/2008-based-projections/index.html

ONS 2012 Family Spending 2011. Table A52

ONS. 2013a. The 2013 ONS regional characteristics analysis for Yorkshire and The Humber [Accessed June 8th 2014]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/rel/regional-trends/region-and-country-profiles/region-and-country-profiles---key-statistics-and-profiles--october-2013/key-statistics-and-profiles---yorkshire-and-the-humber--october-2013.html

ONS. 2013b. Region and Country Profile: Directory of Tables. [Accessed June 4th 2014]. Available from http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171766_344851.pdf

ONS. 2014. E-commerce: measuring, monitoring and gross domestic product. Available from: http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/articles/2011-present/e-commerce--measuring--monitoring-and-gross-domestic-product.pdf.

ONS. 2015a. E-Commerce and ICT Activity: 2014. [Accessed July 1st 2016]. Available from: http://www.ons.gov.uk/businessindustryandtrade/itandInternetindustry/bulletins/ecommerceandictactivity/2014#e-commerce-sales-by-geographical-area

ONS. 2015b. Overview of Internet Sales 2014. [Accessed June 2nd2016]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/rel/rsi/retail-sales/december-2014/sty-overview-of-Internet-retail-sales-in-2014.html

ONS. 2015c. Internet Users – 2015. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/dcp171778_404497.pdf

ONS. 2016. The relationship between gross value added (GVA) and gross domestic product (GDP). [Accessed August 17th 2016]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/gva/relationship-gva-and-gdp/gross-value-added-and-gross-domestic-product.html

ONS. 2016a. Internet access-Households and Individuals 2016. [Accessed June 5th 2016]. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeInternetandsocialmediausage/bulletins/Internetaccesshouseholdsandindividuals/2016

ONS. 2016b. 2011 OAC Area Classifications. [Accessed June 6th 2016]. Available from: http://webarchive.nationalarchives.gov.uk/20160105160709/http://ons.gov.uk/ons/guide-method/geography/products/area-classifications/national-statistics-area-classifications/national-statistics-2011-area-classifications/index.html

Openshaw, S. 1976. An Empirical Study of Some Spatial Interaction Models. *Environment and Planning A*, 8: pp. 23-41.

Palmer, N. 2010. *UK Commerce figures*. IM3 Marketing. 15 August 2010. [Accessed January 31st 2013]. Available from: http://www.im3.co.uk/e-commerce/uk-ecommerce-figures/

Palmer, N. 2012. *E-Commerce Secures Future for UK Business growth*. IM3 Marketing. 12 July 2012. [Accessed February 1st 2013]. Available from: http://www.im3.co.uk/e-commerce/e-commerce-secures-future-for-uk-business-growth/

Reynolds, J., Wood, S. 2010. Location Decision Making in Retail Firms: Evolution and Challenge. [Accessed July 7th 2014]. Available from: http://epubs.surrey.ac.uk/209759/1/Location%20Decision-Making%20in%20Retail%20Firms%20Evolution%20and%20Challenge.pdf

Retail Gazette. 2016. £114bn spent online in 2015. [Accessed February 15th 2016]. Available from: http://www.retailgazette.co.uk/blog/2016/01/ps114bn-spent-online-in-2015

Riddlesden, D., Singleton A. 2014. Broadband speed equity: A new digital divide? *Applied geography.* 52 (8), pp. 25-33.

Roger E. 2003. *The Diffusion of Innovations*. 5th ed.The Free Press. New York

Ruddick G. 2014. *It may already be too late for Tesco and Sainsbury's, the rise of Aldi and Lidl looks unstoppable* [Accessed May 14th 2015]. Available from:
http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/

Salomon, I. and Koppelman, F., 1988. A framework for studying teleshopping versus store shopping. *Transportation Research Part A: General*, *22*(4), pp.247-255.

Seely, A. 2012. Supermarkets : competition inquiries into the groceries market. *House of Commons Library.* [Accessed May 14th 2013]. Available from: www.parliament.uk/briefing-papers/sn03653.pdf

Sexton, R., Johnson, R. and Hignite, M. 2002. Predicting Internet/E-Commerce Use. Internet
Research. *Electronic Networking and Policy*, 12(5), pp. 402-410.

Simkin, L., Doyle, P. & Saunders, J. 1985. How retailers put site location techniques into operation: An assessment of major multiples' practice Retail & Distribution Management, 13(3), pp. 21-26.

Simmons J.W. 1978. Systems of cities: readings on structure growth and policy.

Singh, N., Lehnert, K., Bostick, K. 2012. Global social media usage: insights into reaching consumers worldwide. *Thunderbird International Business Review*, 54, pp. 683-700.

Singleton A. 2014. *The Past, Present, and Future of Geodemographic Research in the United States and United Kingdom* [Accessed May 14th 2016]. Available from: http://www.tandfonline.com/doi/full/10.1080/00330124.2013.848764?src=recsys

Statista. 2016a. B2C e-commerce as percentage of global GDP from 2009 to 2018. [Accessed June 1st 2016]. Available from: http://www.statista.com/statistics/324612/b2c-e-commerce-s-percentage-of-gdp-worldwide/

Statista. 2016b. Share of households with Internet access in the United Kingdom (UK) from 1998 to 2015 [Accessed June 6th 2016]. Available from: http://www.statista.com/statistics/275999/household-Internet-penetration-in-great-britain/

Startupfactory, 2014. Understanding and Using Magic Quadrant-style Analysis [Accessed June 6th 2016]. Available from: http://www.startupfactory.co/pdf/3- SUF_magicquads.pdf

The National Farm Research Unit. 2014. Defra's Food Statistics Poketbook 2014. [Accessed March 14th 2015]. Available from: http://www.nfru.co.uk/defra-monthly-briefing.html

The Tinder Foundation. 2016. Social Housing and Digital Inclusion : People not Technology. [Accessed May 7th 2016]. Available from: http://www.tinderfoundation.org/our-thinking/blog/social-housing-and-digital-inclusion-people-not-technology

The Financial Times. 2016. Delivery charges cost online retailers dear. [Accessed June 4th 2016]. Available from: http://www.ft.com/cms/s/2/fd88f556-70bc-11e5-9b9e-690fdae72044.html#axzz4GGtx1bJv

The Guardian. 2015. The UK uses the most tracking cookies of any EU country. How should you be protecting your privacy online? [Accessed July 1st 2016]. Available from: https://www.theguardian.com/technology/2015/mar/19/cookies-how-to-avoid-being-tracked-online

The Guardian. 2016. As Amazon takes on the UK grocery market, can it delver a profit? [Accessed July 1st 2016]. Available from: https://www.theguardian.com/technology/2016/jun/11/amazon-fresh-launches-in-uk-can-it-deliver

The Telegraph. 2016. The four reasons why supermarkets are losing money. [Accessed June 10th 2015]. Available from: http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11585366/The-four-reasons-why-supermarkets-are-losing-money.htm

The Telegraph. 2016. Online shopping to grow by £320bn in three years. [Accessed July 1st 2016]. Available http://www.telegraph.co.uk/finance/newsbysector/retailandconsumer/11657830/online-shopping-to-grow-by-320bn

Thomas, R.W., Hugget, D.J. 1980. Modelling in Geography: A Mathematical Approach. Harper and Row, London.

Thompson, C. 2012. Retailers Average Floorspace [Personal Communication]. *Personal Communication*:25th March 2011. Leeds.

Tradeglobal. 2015. Bridging the gap between social media and e-commerce. [Accessed May 6th 2016]. Available from: http://www.tradeglobal.com/bridging-gap-social-media-ecommerce/

Vickers, D., Rees, P., Birkin, M. 2005. Creating the National Classification of Census Output Areas: Data, Methods and Results. [Accessed March 7th 2013]. Available from: http://www.geog.leeds.ac.uk/fileadmin/documents/research/csap/05-02.pdf

Walters, D. 1974. Retail site location: time for a new approach? *Retail & Distribution Management,* 2(6), pp. 28- 31.

Ward and Morganovsky, 2002

Watershed Publishing. 2016. Internet Economy Reported Contributing to 4.7% of US '10 GDP. [Accessed March 12th 2015]. Available from: http://www.marketingcharts.com/uncategorized/Internet-economy-found-contributing-to-almost-5-of-us-gdp-in-10-21549/

Webinterpret. 2015. United Kingdom: top online market and eCommerce success story. [Accessed March 3rd 2016]. Available from: http://webinterpret.com/blog/cross-border-trade-in-practice/united-kingdom-top-online-market-e-commerce-success-story/

Weltevreden, J. 2006. Cyberspace meets high street: Adoption of click-and-mortar strategies by retail outlets in city centres. *Urban Geography*, *27*(7), pp.628-650.

Williams, D. E. 2009. The evolution of e-tailing, International Review of Retail, Distribution & Consumer Research, Vol. 19. Pp. 219-49.

Wilson, A. G. 1971. A family of spatial interaction models, and associated developments. *Environment and Planning* 3, pp.1-32.

Wrigley, N. 1991. Is the 'golden age' of British grocery retailing at a watershed? *Environment and Planning A,* 23. Pp.1537-1544.

Wrigley, N. & Lowe, M. 2002.Reading retail: a geographical perspective on retailing and consumption spaces. GB: Hodder Education. [Accessed October 1st 2012]. Available from: http://www.dawsonera.com/depp/reader/protected/external/AbstractView/S9781444118759

Wrigley, N. and Dolega, L., 2011. Resilience, fragility, and adaptation: new evidence on the performance of UK high streets during global economic crisis and its policy implications. *Environment and Planning A*, *43*(10), pp.2337-2363.

Xing, Y. and Grant, D. 2006. Developing a framework for measuring physical distribution service quality of multi-channel and 'Pure Player' Internet retailers. *International Journal of Retail & Distribution Management*, Vol. 34 Nos 4/5, pp. 278-89. [Accessed June 3rd 2015]. Available from: http://www.emeraldinsight.com/doi/abs/10.1108/09590550610660233

Zijlmans, O. 2010.The changing demand for retail property under the influence of online shopping: a study on the future strategies retailers. Eindhoven: University of Technology. [Accessed July 17th 2016].

Appendix A

ACORN population classification

Acorn Type	Description	Acorn Group	Description	Acorn Category	Description
1	Exclusive enclaves	A	Lavish Lifestyles	1	Affluent Achievers
2	Metropolitan money	A	Lavish Lifestyles	1	Affluent Achievers
3	Large house luxury	A	Lavish Lifestyles	1	Affluent Achievers
4	Asset rich families	B	Executive Wealth	1	Affluent Achievers
5	Wealthy countryside commuters	B	Executive Wealth	1	Affluent Achievers
6	Financially comfortable families	B	Executive Wealth	1	Affluent Achievers
7	Affluent professionals	B	Executive Wealth	1	Affluent Achievers
8	Prosperous suburban families	B	Executive Wealth	1	Affluent Achievers
9	Well-off edge of towners	B	Executive Wealth	1	Affluent Achievers
10	Better-off villagers	C	Mature Money	1	Affluent Achievers
11	Settled suburbia, older people	C	Mature Money	1	Affluent Achievers
12	Retired and empty nesters	C	Mature Money	1	Affluent Achievers
13	Upmarket downsizers	C	Mature Money	1	Affluent Achievers
14	Townhouse cosmopolitans	D	City Sophisticates	2	Rising Prosperity
15	Younger professionals in smaller flats	D	City Sophisticates	2	Rising Prosperity
16	Metropolitan professionals	D	City Sophisticates	2	Rising Prosperity
17	Socialising young renters	D	City Sophisticates	2	Rising Prosperity
18	Career driven young families	E	Career Climbers	2	Rising Prosperity
19	First time buyers in small, modern homes	E	Career Climbers	2	Rising Prosperity
20	Mixed metropolitan areas	E	Career Climbers	2	Rising Prosperity

21	Farms and cottages	F	Countryside Communities	3	Comfortable Communities
22	Larger families in rural areas	F	Countryside Communities	3	Comfortable Communities
23	Owner occupiers in small towns and villages	F	Countryside Communities	3	Comfortable Communities
24	Comfortably-off families in modern housing	G	Successful Suburbs	3	Comfortable Communities
25	Larger family homes, multi-ethnic areas	G	Successful Suburbs	3	Comfortable Communities
26	Semi-professional families, owner occupied neighbourhoods	G	Successful Suburbs	3	Comfortable Communities
27	Suburban semis, conventional attitudes	H	Steady Neighbourhoods	3	Comfortable Communities
28	Owner occupied terraces, average income	H	Steady Neighbourhoods	3	Comfortable Communities
29	Established suburbs, older families	H	Steady Neighbourhoods	3	Comfortable Communities
30	Older people, neat and tidy neighbourhoods	I	Comfortable Seniors	3	Comfortable Communities
31	Elderly singles in purpose-built accommodation	I	Comfortable Seniors	3	Comfortable Communities
32	Educated families in terraces, young children	J	Starting Out	3	Comfortable Communities
33	Smaller houses and starter homes	J	Starting Out	3	Comfortable Communities
34	Student flats and halls of residence	K	Student Life	4	Financially Stretched
35	Term-time terraces	K	Student Life	4	Financially Stretched
36	Educated young people in flats and tenements	K	Student Life	4	Financially Stretched
37	Low cost flats in suburban areas	L	Modest Means	4	Financially Stretched
38	Semi-skilled workers in traditional neighbourhoods	L	Modest Means	4	Financially Stretched
39	Fading owner occupied terraces	L	Modest Means	4	Financially Stretched
40	High occupancy terraces, many Asian families	L	Modest Means	4	Financially Stretched

41	Labouring semi-rural estates	M	Striving Families	4	Financially Stretched
42	Struggling young families in post-war terraces	M	Striving Families	4	Financially Stretched
43	Families in right-to-buy estates	M	Striving Families	4	Financially Stretched
44	Post-war estates, limited means	M	Striving Families	4	Financially Stretched
45	Pensioners in social housing, semis and terraces	N	Poorer Pensioners	4	Financially Stretched
46	Elderly people in social rented flats	N	Poorer Pensioners	4	Financially Stretched
47	Low income older people in smaller semis	N	Poorer Pensioners	4	Financially Stretched
48	Pensioners and singles in social rented flats	N	Poorer Pensioners	4	Financially Stretched
49	Young families in low cost private flats	O	Young Hardship	5	Urban Adversity
50	Struggling younger people in mixed tenure	O	Young Hardship	5	Urban Adversity
51	Young people in small, low cost terraces	O	Young Hardship	5	Urban Adversity
52	Poorer families, many children, terraced housing	P	Struggling Estates	5	Urban Adversity
53	Low income terraces	P	Struggling Estates	5	Urban Adversity
54	Multi-ethnic, purpose-built estates	P	Struggling Estates	5	Urban Adversity
55	Deprived and ethnically diverse in flats	P	Struggling Estates	5	Urban Adversity
56	Low income large families in social rented semis	P	Struggling Estates	5	Urban Adversity
57	Social rented flats, families and single parents	Q	Difficult Circumstances	5	Urban Adversity
58	Singles and young families, some receiving benefits	Q	Difficult Circumstances	5	Urban Adversity
59	Deprived areas and high-rise flats	Q	Difficult Circumstances	5	Urban Adversity
60	Active communal population	R	Not Private Households	6	Not Private Households

61	Inactive communal population	R	Not Private Households	6	Not Private Households
62	Business areas without resident population	R	Not Private Households	6	Not Private Households

Appendix B

COICOP Expenditure Classification (CACI, 2013)

Food and non-alcoholic beverages	Food	CV0111_11T	Rice
Food and non-alcoholic beverages	Food	CV0111_21T	Bread
Food and non-alcoholic beverages	Food	CV0111_22T	Buns, crispbread and biscuits
Food and non-alcoholic beverages	Food	CV0111_31T	Pasta products
Food and non-alcoholic beverages	Food	CV0111_41T	Cakes and puddings
Food and non-alcoholic beverages	Food	CV0111_42T	Pastry (savoury)
Food and non-alcoholic beverages	Food	CV0111_51T	Other breads and cereals
Food and non-alcoholic beverages	Food	CV0112_11T	Beef (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_21T	Pork (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_31T	Lamb (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_41T	Poultry (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_51T	Sausages
Food and non-alcoholic beverages	Food	CV0112_52T	Bacon and ham
Food and non-alcoholic beverages	Food	CV0112_53T	Offal, pate etc
Food and non-alcoholic beverages	Food	CV0112_61T	Other preserved or processed meat
Food and non-alcoholic beverages	Food	CV0112_71T	Other fresh, chilled or frozen edible meat
Food and non-alcoholic beverages	Food	CV0113_11T	Fish (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0113_21T	Seafood (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0113_31T	Dried, smoked or salted fish & seafood

Food and non-alcoholic beverages	Food	CV0113_41T	Other preserved or processed fish and seafood
Food and non-alcoholic beverages	Food	CV0114_11T	Whole milk
Food and non-alcoholic beverages	Food	CV0114_21T	Low fat milk
Food and non-alcoholic beverages	Food	CV0114_31T	Preserved milk
Food and non-alcoholic beverages	Food	CV0114_41T	Yoghurt
Food and non-alcoholic beverages	Food	CV0114_51T	Cheese and curd
Food and non-alcoholic beverages	Food	CV0114_61T	Other milk products
Food and non-alcoholic beverages	Food	CV0114_71T	Eggs
Food and non-alcoholic beverages	Food	CV0115_11T	Butter
Food and non-alcoholic beverages	Food	CV0115_21T	Margarine and other vegetable fats
Food and non-alcoholic beverages	Food	CV0115_22T	Peanut butter
Food and non-alcoholic beverages	Food	CV0115_31T	Olive oil
Food and non-alcoholic beverages	Food	CV0115_41T	Edible oils
Food and non-alcoholic beverages	Food	CV0115_51T	Other edible animal fats
Food and non-alcoholic beverages	Food	CV0116_11T	Citrus fruits (fresh)
Food and non-alcoholic beverages	Food	CV0116_21T	Bananas (fresh)
Food and non-alcoholic beverages	Food	CV0116_31T	Apples (fresh)
Food and non-alcoholic beverages	Food	CV0116_41T	Pears (fresh)
Food and non-alcoholic beverages	Food	CV0116_51T	Stone fruits (fresh)

Food and non-alcoholic beverages	Food	CV0116_61T	Berries (fresh)
Food and non-alcoholic beverages	Food	CV0116_71T	Other fresh, chilled or frozen fruits
Food and non-alcoholic beverages	Food	CV0116_81T	Dried fruit and nuts
Food and non-alcoholic beverages	Food	CV0116_91T	Preserved fruit and fruit based products
Food and non-alcoholic beverages	Food	CV0117_11T	Leaf and stem vegetables (fresh or chilled)
Food and non-alcoholic beverages	Food	CV0117_21T	Cabbages (fresh or chilled)
Food and non-alcoholic beverages	Food	CV0117_31T	Veg grown for their fruit (fresh, chilled, frozen)
Food and non-alcoholic beverages	Food	CV0117_41T	Root crops and mushrooms (fresh, chilled, frozen)
Food and non-alcoholic beverages	Food	CV0117_51T	Dried vegetables
Food and non-alcoholic beverages	Food	CV0117_61T	Other preserved or processed vegetables
Food and non-alcoholic beverages	Food	CV0117_71T	Potatoes
Food and non-alcoholic beverages	Food	CV0117_81T	Other tubers and products of tuber vegetables
Food and non-alcoholic beverages	Food	CV0118_11T	Sugar
Food and non-alcoholic beverages	Food	CV0118_21T	Jams, marmalades
Food and non-alcoholic beverages	Food	CV0118_31T	Chocolate
Food and non-alcoholic beverages	Food	CV0118_41T	Confectionery products
Food and non-alcoholic beverages	Food	CV0118_51T	Edible ices and ice cream
Food and non-alcoholic beverages	Food	CV0118_61T	Other sugar products
Food and non-alcoholic beverages	Food	CV0119_11T	Sauces, condiments

Food and non-alcoholic beverages	Food	CV0119_21T	Salt, spices and culinary herbs
Food and non-alcoholic beverages	Food	CV0119_31T	Bakers yeast, dessert preparations, soups
Food and non-alcoholic beverages	Food	CV0119_41T	Other food products
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_11T	Coffee
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_21T	Tea
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_31T	Cocoa and powdered chocolate
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_11T	Mineral or spring waters
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_21T	Soft drinks
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_31T	Fruit juices
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_41T	Vegetable juices
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0211_11T	Spirits and liqueurs (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_11T	Wine from grape or other fruit (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_12T	Fortified wine (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_13T	Ciders and Perry (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_14T	Alcopops (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_21T	Champagne and sparkling wines (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0213_11T	Beer and lager (brought home)
Food and non-alcoholic beverages	Food	CV0111_11T	Rice
Food and non-alcoholic beverages	Food	CV0111_21T	Bread

Food and non-alcoholic beverages	Food	CV0111_22T	Buns, crispbread and biscuits
Food and non-alcoholic beverages	Food	CV0111_31T	Pasta products
Food and non-alcoholic beverages	Food	CV0111_41T	Cakes and puddings
Food and non-alcoholic beverages	Food	CV0111_42T	Pastry (savoury)
Food and non-alcoholic beverages	Food	CV0111_51T	Other breads and cereals
Food and non-alcoholic beverages	Food	CV0112_11T	Beef (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_21T	Pork (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_31T	Lamb (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_41T	Poultry (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0112_51T	Sausages
Food and non-alcoholic beverages	Food	CV0112_52T	Bacon and ham
Food and non-alcoholic beverages	Food	CV0112_53T	Offal, pate etc
Food and non-alcoholic beverages	Food	CV0112_61T	Other preserved or processed meat
Food and non-alcoholic beverages	Food	CV0112_71T	Other fresh, chilled or frozen edible meat
Food and non-alcoholic beverages	Food	CV0113_11T	Fish (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0113_21T	Seafood (fresh, chilled or frozen)
Food and non-alcoholic beverages	Food	CV0113_31T	Dried, smoked or salted fish & seafood
Food and non-alcoholic beverages	Food	CV0113_41T	Other preserved or processed fish and seafood
Food and non-alcoholic beverages	Food	CV0114_11T	Whole milk

Food and non-alcoholic beverages	Food	CV0114_21T	Low fat milk
Food and non-alcoholic beverages	Food	CV0114_31T	Preserved milk
Food and non-alcoholic beverages	Food	CV0114_41T	Yoghurt
Food and non-alcoholic beverages	Food	CV0114_51T	Cheese and curd
Food and non-alcoholic beverages	Food	CV0114_61T	Other milk products
Food and non-alcoholic beverages	Food	CV0114_71T	Eggs
Food and non-alcoholic beverages	Food	CV0115_11T	Butter
Food and non-alcoholic beverages	Food	CV0115_21T	Margarine and other vegetable fats
Food and non-alcoholic beverages	Food	CV0115_22T	Peanut butter
Food and non-alcoholic beverages	Food	CV0115_31T	Olive oil
Food and non-alcoholic beverages	Food	CV0115_41T	Edible oils
Food and non-alcoholic beverages	Food	CV0115_51T	Other edible animal fats
Food and non-alcoholic beverages	Food	CV0116_11T	Citrus fruits (fresh)
Food and non-alcoholic beverages	Food	CV0116_21T	Bananas (fresh)
Food and non-alcoholic beverages	Food	CV0116_31T	Apples (fresh)
Food and non-alcoholic beverages	Food	CV0116_41T	Pears (fresh)
Food and non-alcoholic beverages	Food	CV0116_51T	Stone fruits (fresh)
Food and non-alcoholic beverages	Food	CV0116_61T	Berries (fresh)
Food and non-alcoholic beverages	Food	CV0116_71T	Other fresh, chilled or frozen fruits

Food and non-alcoholic beverages	Food	CV0116_81T	Dried fruit and nuts
Food and non-alcoholic beverages	Food	CV0116_91T	Preserved fruit and fruit based products
Food and non-alcoholic beverages	Food	CV0117_11T	Leaf and stem vegetables (fresh or chilled)
Food and non-alcoholic beverages	Food	CV0117_21T	Cabbages (fresh or chilled)
Food and non-alcoholic beverages	Food	CV0117_31T	Veg grown for their fruit (fresh, chilled, frozen)
Food and non-alcoholic beverages	Food	CV0117_41T	Root crops and mushrooms (fresh, chilled, frozen)
Food and non-alcoholic beverages	Food	CV0117_51T	Dried vegetables
Food and non-alcoholic beverages	Food	CV0117_61T	Other preserved or processed vegetables
Food and non-alcoholic beverages	Food	CV0117_71T	Potatoes
Food and non-alcoholic beverages	Food	CV0117_81T	Other tubers and products of tuber vegetables
Food and non-alcoholic beverages	Food	CV0118_11T	Sugar
Food and non-alcoholic beverages	Food	CV0118_21T	Jams, marmalades
Food and non-alcoholic beverages	Food	CV0118_31T	Chocolate
Food and non-alcoholic beverages	Food	CV0118_41T	Confectionery products
Food and non-alcoholic beverages	Food	CV0118_51T	Edible ices and ice cream
Food and non-alcoholic beverages	Food	CV0118_61T	Other sugar products
Food and non-alcoholic beverages	Food	CV0119_11T	Sauces, condiments
Food and non-alcoholic beverages	Food	CV0119_21T	Salt, spices and culinary herbs
Food and non-alcoholic beverages	Food	CV0119_31T	Bakers yeast, dessert preparations, soups

Food and non-alcoholic beverages	Food	CV0119_41T	Other food products
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_11T	Coffee
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_21T	Tea
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0121_31T	Cocoa and powdered chocolate
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_11T	Mineral or spring waters
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_21T	Soft drinks
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_31T	Fruit juices
Food and non-alcoholic beverages	Non-alcoholic beverages	CV0122_41T	Vegetable juices
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0211_11T	Spirits and liqueurs (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_11T	Wine from grape or other fruit (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_12T	Fortified wine (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_13T	Ciders and Perry (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_14T	Alcopops (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0212_21T	Champagne and sparkling wines (brought home)
Alcoholic beverages and tobacco	Alcoholic beverages (off-sales)	CV0213_11T	Beer and lager (brought home)