Consumer Preferences & Loyalty Discounts
The Case of UK Grocery Retail

Domile Butkeviciute

PhD Economics

City, University of London

Department of Economics

November 2017

Table of Contents

List of Tables

List of Figures:

Chapter I

Chapter II

Chapter III

Appendix

Dedication

To my incredible grandparents who will always be my greatest inspirations.

Skiriu savo Močiutėms ir Diedukui, kurie man visados išliks didžiausias įkvėpimo šaltinis.

Acknowledgements

I extend my gratitude to the Department of Economics, faculty members, my fellow PhD candidates and the brilliant students who I had the privilege of teaching. Thank you for challenging me, sharing your knowledge, experience and good humour. A special thank you goes to my thesis supervisor Dr Xeni Dassiou for her encouragement and support. I would also like to thank Professors Michael Ben-Gad and Jon Stern for backing me and providing me with invaluable advice during my time as a research student and in the context of my teaching career.

To all my friends, thank you for your kindness, patience and for always cheering me on and believing in me. From the bottom of my heart, thank you to Mark for helping me overcome the challenges I faced, for persevering with me throughout this rollercoaster journey and helping me see the light at the end of the tunnel. Of course, none of my achievements would be a reality without my inspirational parents. Thank you both, for helping me reach for the highest stars and throughout all my ups and downs, the tears and the sweat, you loved me, motivated me, pushed me and supported me beyond the imaginable.

Declaration

I hereby grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

This thesis re-examines a common assumption entering theoretical models of endogenous switching costs. Through a discrete choice experiment we test the hypothesis that consumers are heterogeneous in the way they respond when firms offer repeat purchase discounts through loyalty schemes. The assumption itself is important because in practice, heterogeneity in consumer switching costs holds implications for firms' strategies and their resultant market shares. This thesis presents a flexible methodology for a discrete choice experiment inspired by the UK groceries sector using novel techniques in *D*-efficient experimental survey design. When fitting the data to the mixed logit model, we find that consumers' taste varies significantly more for loyalty schemes than for any of the other variables entering the model. The results of our discrete choice experiment show that consumers differ significantly in how they respond to repeat purchase discount strategies. On this basis, it is likely that theoretical models of loyalty schemes overemphasise the effects of loyalty schemes on price competition. We argue instead, that a repeat purchase discount strategy will not result in a unilateral increase in artificial switching costs for all consumers in the market. We propose that forward looking firms are likely to recognise the limitations to scheme effectiveness due to heterogeneity in switching costs and will be more likely to invest in their customer base through future lower prices. Therefore from a competition policy perspective, we argue that in a fast paced retail market for non-durable goods, loyalty schemes are more likely to intensify competition for the benefit of consumers rather than act as an exclusionary device.

Thesis Introduction

Traditionally, models of endogenous switching costs have assumed that consumers are homogeneous in the way they perceive artificial costs of switching when firms offer repeat purchase discounts. This thesis re-examines this assumption and shows empirically that consumers' responses to repeat purchase discount strategies are likely to vary significantly in real-world markets. This result is driven by the fact that consumers are *heterogeneous* in their costs of switching when retailers offer loyalty schemes. It follows that *some* consumers do *not* perceive any switching costs when choosing between retailers when loyalty schemes are introduced in the market. Instead, these consumers base their choice of retailer on their general preferences and this is independent of the availability of a repeat purchase discount. In fact, we find that a loyalty scheme may actually *reduce* utility for some individuals who may for example, prefer not to allow retailers to collect, store and analyse their personal data. Therefore, on the basis of our empirical findings, we argue that the theory cannot simply assume that repeat purchase discounts unilaterally increase switching costs for all consumers. In turn, we note the limitations of scheme effectiveness will likely affect the strategies adopted by forward looking firms competing for consumers in a non-durable goods market.

From a competition policy perspective, switching costs are generally viewed as welfare *reducing* because as consumers become "locked-in" through perceived and/ or actual costs of switching they are less willing to switch away in future periods of competition. This demand side effect has a tendency to soften price competition. In this context, forward looking firms may choose to create switching costs by implementing loyalty schemes strategies. Considering the industrial organisation literature, it is well-documented that consumers incur costs of switching between brands and providers in many markets (Klemperer 1995, Farrell & Klemperer 2007). For example, individuals' psychological attachments to brands are associated with switching costs and may reduce their willingness to try alternative products. Transaction costs incurred by customers when switching between providers also represent a type of switching cost.

Both of the above examples represent *exogenous* switching costs and prevail in markets regardless of consumers' behaviour or firms' strategies. On the other hand, *endogenous* switching costs arise when firms make strategic decisions which create artificial switching costs. Namely by entering into contractual arrangements with customers, introducing product incompatibility or offering customers repeat purchase discounts in the form of loyalty schemes. Exogenous and endogenous switching costs are generally associated with less competitive markets and tend to soften price competition (Klemperer 1995, Farrell & Klemperer 2007). However, they may also lead firms to compete more fiercely for market share through lower prices (Rhodes 2014). There are a number of drivers underlying these outcomes and we briefly consider these below.

It has been shown that in markets with exogenous switching costs firms compete vigorously *ex ante* to gain and establish a large market share to achieve greater market power in future periods over locked-in consumers (Farrell & Klemperer 2007). With the knowledge that consumers are partially locked-in in a market with exogenous switching costs, firms tend to face the trade-off between *harvesting* or *investing* in market share (Klemperer 1995, Anderson & Kumar 2007, Rhodes 2014). In some cases this leads to a bargain-to-rip-off game where firms offer low prices to grow market share and then charge higher prices in future periods (Farrell & Klemperer 2007). In other words, anticipating that consumers will not switch away, a firm with a degree of market power can harvest customers by charging higher prices. Alternatively, a firm may instead invest in growing its market share and reduce prices leading to more intense price competition. Therefore the assumptions entering a theoretical model, namely the number of stages of competition, will have a direct impact on the competitive effects of switching costs and market outcomes (Rhodes 2014).

Equally, loyalty scheme strategies can lead to either *more* or *less* competitive markets depending on consumers' responses, the market structure and firm symmetry. For instance, loyalty schemes are generally considered to be anticompetitive in duopoly, while under monopolistic competition, loyalty schemes are associated with largely procompetitive effects that enhance welfare for consumers (Caminal & Claici 2007). The assumptions on consumer preferences are also important in such models as loyalty scheme effectiveness drives consumer responses and the firm's own strategy going forward. However, theoretical models of repeat purchase discounts, by definition assume that firms themselves determine the size of artificial switching costs (Klemperer 1995, Farrell & Klemperer). Another important assumption found in theoretical models of switching costs, is that consumers incur switching costs *homogenously* say through a unilateral increase in transport costs in a Hotelling framework. As far as we are aware, there exists only one recent publication which considers heterogeneous switching costs (Biglaiser et al. 2016). The authors provide an important contribution in that "[…] *heterogeneity of switching costs has complex strategic consequences which have largely been ignored in the literature. It will influence the strategies of firms, the equilibrium distribution of clients, and the value of incumbency*."[1]

This thesis is structured as follows. Section 1.2 of the first chapter sets out the features of UK groceries sector which represents a mature market for loyalty scheme strategies. We then set out in detail the main aspects of the theoretical literature on switching costs and loyalty schemes in nondurable goods markets in Section 1.3. Section 1.4 then presents empirical evidence on the effects of loyalty schemes. In reviewing the literature, we find that few papers study the demand-side effects of the strategy in fast moving retail markets and attention has instead been typically placed on frequent flier programs. The effects of the strategy are also influenced by the principal agent problem which is not common across markets. In the context of the review, we identify a gap in the literature in the context of heterogeneity in endogenous switching costs when firms implement repeat purchase discounts in fast moving consumer markets.

Section 1.5 sets out the relevant competition policy considerations for markets with endogenous and exogenous switching costs. In doing so we apply the lessons from the theory and also assess the relevant practical considerations for intervention. This section looks at the recent market investigations into UK retail banking and the retail energy where low switching by consumers was deemed as being particularly problematic for rivalry between firms and outcomes for consumers. Towards the end of the first chapter, we introduce the concept of discrete choice experiments. This approach enables the researcher to address very specific questions on drivers of consumer choice by mimicking real-world markets through some form of instrument. Either through a series of survey questions or controlled lab experiments. The approach also overcomes some of the limitations of theoretical models which may not capture the wider aspects of a retail offer and instead assume a relatively unsophisticated consumer preferences. We do acknowledge that theoretical modelling is essential in explaining broader dynamics of markets to help us understand the underlying rationale

[1] Biglaiser, G., Crémer J., Dobos, G., 'Heterogeneous switching costs', *International Journal of Industrial Organization*, Vol. 47, p. 63, 2016.

and incentives for the pricing and discounting strategies adopted by firms. However, a typical model focuses on the firm and rivalry with competitors *not* the consumer. This is the case even though consumers' preferences and reactions to strategies can have significant effects on the competitiveness of a market. We recognize that DCEs are useful *however* these are of course experiments. As such, they will always be imperfect due to the complex nature of real world markets (Waterson 2014). Nonetheless, we argue that when analysing the effectiveness of business strategies which rely heavily on reactions of consumers, a combination of theoretical and experimental evidence may be optimal. In combining both approaches, we note that empirical evidence can help determine realistic assumptions to enter a theoretical model. The model itself can then be applied for a broader assessment of welfare effects and consumer outcomes more accurately reflecting the demand and supply sides of the market.

The experiment we design in this thesis relies on novel techniques in stated choice methods. The theoretical underpinnings and applications of stated choice methods are presented within the beginning of the second chapter of this thesis in Section 2.2. Section 2.3 derives the functional form of the logistic regression models used to analyse micro-level choice data including the conditional logit model and the more advanced mixed logit. The mixed logit model accommodates preference heterogeneity and allows us to later show that consumers differ in their taste for repeat purchase discounts. In the second chapter we note that there exists two main methods for the design of the unique survey: a traditional orthogonal design or an efficient design. Section 2.4 sets out the trade-offs inherent to these methods noting that our chosen *efficient experimental design* approach requires a smaller sample size to achieve robust parameter estimates. This methodology incorporates the discrete choice model into the survey design itself and thus seeks to reduce the resulting standard errors of parameter estimates. Furthermore, unlike orthogonal designs which are based on the statistical properties of linear models, the data obtained using an efficient survey design can be easily accommodated by non-linear discrete choice models. Section 2.5 sets out additional considerations for the design of a survey in terms of contextual realism, the sample size requirements, how to select the optimal number of questions to present to respondents and other relevant factors of survey design.

Ahead of presenting our detailed methodology, Section 2.6 evaluates ways in which researchers perform surveys of human populations, noting the possible biases tied to online surveys and convenience sampling approaches. Then, we outline a methodology for the design of a discrete choice experiment to measure consumer preferences for loyalty schemes in the UK groceries market in Section 2.7. We explain how we design the survey questions to maximise elicitation of truthful response, in particular on sensitive issues such as income. We also perform a qualitative assessment of this market to ensure the experiment accurately reflects features of the sector. We undertake a small pilot study and evaluate the results in Section 2.8. We rely on the estimated coefficients to set the parameter priors to enter the final design of our survey. This is a requirement in *D*-efficient designs. The final survey design is outlined in Section 2.9 where we explain how to control for uncertainty on parameter prior values through Bayesian estimation methods.

Proceeding to the third chapter, we present the empirical results of the discrete choice experiment. We begin in Section 3.2 by outlining how we cleaned the data by removing certain responses. We also perform a number of data quality checks to ensure that the results we estimate are in fact reasonable. We then compare the sociodemographic characteristics of survey respondents to the general population figures in section 3.3. We find that the sampled respondents are wealthier, younger and on average more likely to be from Southern regions where household incomes are above the national average. This source of bias is addressed within the empirical analysis section through the application of weights in Section 3.4. Section 3.5 presents the results of further model specifications where we perform individual-level parameter estimation as well as presenting estimates of willingness-to-pay. We then apply our empirical results to the literature we reviewed as part of the first chapter as part of the discussion in Section 3.6.

The coefficient estimates we obtain suggest that consumers have heterogeneous preferences for most grocery retailer attributes. However, we observe the most variation in taste for the repeat purchase discount attribute. Our results suggest that around a third of consumers and/ or households prefer not to receive a repeat purchase discount. When looking at the distribution in tastes (i.e. unobservable variation in preferences) among grocery shoppers, the discount variable displays the most variation with 68% of grocery shoppers favouring a loyalty scheme when choosing between grocery retailers. While the remaining 32% of shoppers prefer not to participate in any loyalty scheme at all. On the basis of our empirical results, we argue that some consumer segments are more likely to incur artificial switching costs than others. For example, individuals who participate in *few* loyalty schemes tend to be more affected by the presence of a loyalty scheme. In terms of behaviour, they will be more likely to choose to participate in a loyalty scheme and make their future retailer choices on that basis. In addition, the results show that while the loyalty scheme coefficient is positive for most individuals, this is not the case for everyone. On the basis of our estimates, respondents who indicated that they do not participate in any loyalty schemes at all, may actually receive a *disutility* from this attribute. This suggests that some consumers may even be deterred by the scheme perhaps due to data privacy concerns. As such, these shoppers will instead choose the retailer corresponding to their current preferences, even if it is the retailer who offers the loyalty scheme. In reality the availability of a loyalty scheme does not necessarily imply the consumer is forced to sign up to the scheme and redeem a coupon.

Our results strongly suggest that it is unrealistic to assume that when a firm implements a loyalty rewarding scheme this will unilaterally increase artificial switching costs for all consumers active in the market. As loyalty schemes do *not* create artificial switching costs for all consumers, the effects of the strategy are likely to be *weaker* and thus lead to a less significant impact on price competition than suggested in the literature. In applying these aspects of our results to the theoretical assumptions on consumer behaviour, we argue that variation in endogenous switching costs between consumers may have ambiguous welfare effects depending on the model design. In particular, whether rival firms can actually observe this aspect of differentiation in behaviour and how they would react in response. For example, by price discriminating between groups of consumers.

As such, in a nondurable goods market, absent significant exogenous costs of switching, loyalty schemes will have weaker lock-in effects than suggested by theory and will be more likely to be procompetitive in nature. We argue that due to limitations to scheme effectiveness outlined above, firms likely face weaker incentives to engage in *harvesting* of consumers and instead, are more likely to choose a strategy which seeks to *invest* in market share. Further, assuming rival firms cannot discriminate in their pricing between different groups of consumers who vary in their sensitivities to loyalty schemes, they will be incentivised to implement other strategies to compete for different *types* of consumers, namely through higher quality or lower average prices.

Chapter I

Literature Review

1.1 Introduction

In some retail markets, firms have the profit incentive to implement loyalty scheme strategies that offer consumers repeat purchase discounts (Klemperer 1987, 1995, Caminal & Claici 2007). The strategy creates artificial switching costs and as a result, consumers are less willing to switch away to an alternative product or provider (Klemperer 1995). It has been established in the literature, that switching costs allow firms to enjoy a degree of market power over the segment of "locked-in" consumers giving them a profit incentive to maintain this position. It follows that the lock-in effect created by switching costs, explains to an extent, firms' willingness to invest in increasing and/ or maintaining high current market shares as this will lead to greater profits in future. Following Klemperer (1995), from the perspective of the firm, it chooses to offer repeat purchase discounts if it anticipates that the current cost of implementing the scheme will be outweighed by the profits it can achieve as a result. Considering the above, firms prefer larger switching costs and not having to commit to future prices where they can recoup the discounts offered to customers.

It can be shown that rivalry between firms attempting to secure higher market share in the presence of switching costs can either *soften* or *intensify* competition (Klemperer 1987, Caminal & Matutes 1990, Caminal & Claici 2007). The type of competitive effect depends on a number of factors including the number of firms, firm symmetry, number of periods entering a theoretical model, presence of different types of switching costs among other factors. We note however, that the models used to assess the effects of switching costs typically ignore the fact that consumers are likely to be heterogeneous in their switching costs (Biglaiser et al. 2016). This is the case even if switching costs have significant effects on firms' strategic decisions and distribution of market shares (Biglaiser et al. 2016). Further, theoretical models in industrial organisation typically focus on the behaviour of firms, ignoring important aspects of consumer behaviour (Waterson 2003). This is the case even if in some markets consumers behaviour affects the functioning of markets and incentives of competing firms. In such cases, the effects of consumer behaviour, namely costs of switching, should be reflected in competition policy (Waterson 2003).

This thesis begins in Section 1.2 by looking at features of the UK groceries sector, where two of the largest retailers by market share, Tesco and Sainsbury, offer loyalty schemes that enable consumers to earn and spend accumulated points in a variety of ways.[2] The evidence, albeit limited, on the UK grocery retailer Tesco's Clubcard, indicates that loyalty schemes can be a tool to gain and retain consumers (Rowley 2007, Turner & Wilson 2006). For example, research by The Institute of Grocery Distribution published in May 2013, reports that 43% of shoppers stated that the ability to use a loyalty scheme in store determines their choice of grocery retailer.[3] More recently, high-end grocery retailer Waitrose, also introduced its own loyalty scheme offering a free coffee and newspaper to scheme members on a daily basis. The trend is set to continue with businesses showing an appetite for innovations in mobile payment systems and corresponding mobile apps that accommodate more sophisticated loyalty programs offering additional customer insights.[4]

After outlining the main features of the groceries sector, we present the economic theory of switching cost which explains the rationale for the use loyalty schemes in retail markets. In presenting the literature, we note that despite the growing popularity of loyalty rewarding schemes in a number of markets, there exist only a few publications which focus specifically on the competitive effects of the strategy, particularly from an empirical perspective. Thus the impact of loyalty programs on competition and outcomes is not completely established in neither marketing[5] nor industrial organization research (Caminal & Claici 2007, Caminal 2012, Dorotic, Bijmolt & Verhoef 2012). Most models used to analyse loyalty scheme strategies are derived from the theory of endogenous switching costs.[6] In this setting, firms create switching costs through strategic decisions which may include the adoption of repeat purchase discounts. The resultant market shares of firms depend on rivals' actions, particularly their ability to respond with a lower price or similar discount. Thus it can be shown that even when firms are largely identical *ex ante,* the repeat purchase discount can help the firm differentiate itself from rivals, act as a business stealing device and facilitate exclusion of rivals in future periods (Caminal & Claici 2007).

On the other hand, loyalty schemes can also be shown to intensify competition between firms attempting to attract consumers through lower prices. Thus loyalty schemes are largely pro-competitive when there are a number of competing firms in the

---

[2] See Figure 1.1 below for the market shares of the main grocery retailers in the UK groceries sector.
[3]IGD, 'What impact do loyalty schemes have on store choice?' 15th July 2013
[4] Mobile payments groups set sights on winning over UK wallets, *The Financial Times*, June 2, 2014
[5] We note that this thesis focuses on the insights offered by the economic models in industrial organisation and we place little weight on the marketing literature.
[6] We note that the Lal & Bell (2003) model of loyalty rewarding programs applies an alternative model framework to that observed in models of endogenous switching costs.

market (Caminal & Claici 2007). However, as noted above, the fact that consumers may not react in the same way to the availability of a repeat purchase discount has largely been ignored in the literature. In addition, the models of loyalty schemes we present further below in Section 1.3, do not account for the rich set of features of retailers in real world markets who compete not only on price but across a number of non-price factors. Section 1.4 considers the existing empirical evidence on loyalty schemes and loyalty inducing discounts. We note that the majority of literature is concerned with frequent flier programs which represents a distinctive market with its own idiosyncrasies, namely the principal agent problem in driving prices paid by customers.

After reviewing the main sources of evidence on different types of switching costs which prevail in markets, we look at likely implications for competition policy in Section 1.5. We look at the recent market investigations into the retail energy and retail banking markets in the UK where switching costs were found to be particularly problematic and leading to poorer outcomes for some customers. In doing so we explain how policy should look at the wider features of a market when assessing the effects of switching costs, namely markets shares of firms over time and price trajectory over time. We also note that while there are inherent trade-offs in the approach, discrete choice methods can help test the assumptions entering theoretical models and accommodate modelling of consumers' variation in taste.

1.2 Case Study: UK Groceries Market

In this section we discuss the role of loyalty schemes in the UK groceries sector and compare the features of Tesco's and Sainsbury's loyalty schemes. These two retailers were the first to implement loyalty schemes and have also enjoyed the highest market shares in the groceries sector. Figure 1.1 below displays the market shares and loyalty scheme launch dates of the main players in the market. A survey of 60 Clubcard holder respondents revealed that the majority were satisfied with the returns received and card ownership was correlated with consumer loyalty (Turner & Wilson 2006). Rowley (2007) argues that the success of this particular loyalty scheme has been in part a result of its multi-dimensional reward design and customer focused approach. Overall the Clubcard is highly integrated into the structure of the company and is a key driver of its brand strategy (Rowley 2007). Further, in a groceries market '*even small shifts in buying habits, multiplied by very large numbers of customers, can provide a welcome* *boost to profits*'.[7] Hence, a well implemented loyalty scheme can be an important mechanism for differentiation in a competitive market. We note that the UK's then Competition Commission ("CC") investigated the groceries sector in 2008. However, the investigation did not consider loyalty schemes as part of its competitive assessment of the market.[8]

We note that in addition to loyalty schemes, over the years retailers in the market have adopted a number of competing business models to attract consumers to their stores. Over the years, retailers implemented price-match promises, repeat purchase coupons and notably Lidl and Aldi have emerged as strong competitors in the discount price segment. For example, at one stage, retailers offered customers a bundled petrol discount for exceeding a certain basket price, in the form of a 5p per litre petrol discount.[9] Further, Sainsbury's and Tesco also implemented a scheme that guaranteed customers a coupon at check out when their shopping basket was more expensive than a comparable rivals' basket.[10]

Figure 1.1 – *UK Grocery Retailer Market Shares & Loyalty Schemes*

	TESCO	Sainsbury's	ASDA	Morrisons	Waitrose
Market Share August 2013	30.3%	16.5%	17.2%	11.3%	4.8%
Market Share August 2015	28.3%	16.3%	16.6%	10.8%	5.1%
Loyalty card since	1995	2002	N/A	2014	2011

Source: Market share data sourced from Kantar Worldpanel other information taken from The Economist, The Guardian and retailer websites.

Price guarantees more generally, are intended to signal low prices to customers. However, all retailers adopting similar price guarantees signal the same message to

[7] *Loyalty Rewards and Insurance: Every little Helps*, The Economist, November 2011
[8] The main concerns identified in the report related to supply chain practices (Competition Commission 2008)
[9] Tesco had several promotions over the years for example If you spend over £50 you receive 5p off per litre of petrol
[10] The Guardian, 'Tesco to accept Sainsbury's Brand Match money-off vouchers', 11 April 2016, http://www.theguardian.com/money/2016/apr/11/tesco-to-accept-sainsburys-brand-match-money-off-vouchers.

customers in the market (Hviid 2011). In turn consumers perceive prices to be more or less equivalent across retailers who offer such price match promises. In this context, consumers are likely to choose their preferred grocery retailer by considering other aspects of the retail offer other than price, for instance location. Given the homogenous nature of products in grocery retail, loyalty schemes can be an important tool for grocery retailers wishing to differentiate themselves from rivals. As outlined in the above figure, Tesco and Sainsbury's were the first to offer loyalty reward schemes to their customers. Following the introduction of the Clubcard in 1995, Sainsbury's introduced the Nectar Card in 2002.[11] Tesco therefore had a first mover advantage in this respect. We note that in 2011, the Tesco Clubcard had 15 million subscribers compared to the 18 million subscribed to the Nectar Card.[12] This difference in subscriber numbers can be explained by the fact that the Nectar Card is available at other participating retailers, while customers of Tesco can only use the Tesco Clubcard at Tesco outlets. In the past, Asda experimented with loyalty cards, however ultimately has stuck to the slogan '*No Clubcard. No gimmicks. Just lower prices every day'*. Interestingly, high end retailer Waitrose, was known for criticizing loyalty cards in the press for an invasion of privacy. Nonetheless on October 25th 2011 they introduced 'MyWaitrose'.[13] The MyWaitrose loyalty card follows a different model and offers participants a daily free newspaper and coffee.

Tesco offers its consumers a fixed ratio loyalty program, where the monetary value of the loyalty discount depends on a customer's total purchases over a given time period. For every pound spent, the customer receives 1 Clubcard point. [14] Once 100 Clubcard points have been earned, the customer earns the equivalent of a £1 voucher. Thus the more products that a customer purchases at Tesco, the higher the reward they receive. Tesco's customers receive a voucher based on the value of accumulated points on a quarterly basis, either electronically or physically by receiving paper vouchers in the post. These vouchers can be saved over the year or even years to be spent by customers on bigger value rewards, like a holiday, for example. Similarly to the Tesco

[11] *Retailing: Spies in your Wallet,* The Economist Printed Issue, November 5 2011
[12] *Ibid.*
[13] *Ibid.*
[14] This promotion of £1 = 2 pts was running since 2009 but is no longer in place.

Clubcard, Sainsbury's offer 2 points per £1 spent, but consumers are rewarded with a £1 coupon for every 200 points accumulated.[15]

Participants of both the Tesco Clubcard and Sainsbury's Nectar Card, can spend their accumulated loyalty points in the form of coupons or e-coupons. Rewards can be spent on an extensive list of leisure activities including an airline discount or hotel voucher.[16] Retailers often engage in further promotional activities such as 'double your points' or 'quadruple your points' for select activities and/or products. These vouchers can then be spent in store or on products and services provided by other companies in different markets namely, entertainment and travel. Such promotions offer consumers an added incentive to spend the loyalty points on a reward. In turn, the cost of the coupon to the retailer is likely to be, at least partially, internalized by another company who accepts the coupons through an agreement.

An important distinction between the two leading retailers is that Tesco is a highly integrated firm and likely relies on fewer agreements to enable consumers to collect points across different products. For example, Tesco operates its own bank and insurance services and therefore enjoys a greater flexibility in the offers it can make to customers. This advantage is not available to Sainsbury, instead is bound by agreements with insurers. Further, Sainsbury's credit card it is operated by American Express. In the past, Sainsbury allowed customers to collect Nectar points when purchasing insurance through a specific price comparison website. On the other hand, Clubcard points can be collected by purchasing products sold by Tesco (including insurance products) and/ or alternatively, points can be collected by using the Tesco credit card on purchases of any item.

Loyalty programs, such as the Tesco Clubcard, confer benefits to consumers but are also a mechanism used by retailers to gather information on consumer behaviour and to help tailor their offers. Tesco is active across a variety of different grocery retail formats and is present in other markets. Tesco Home stores for example are aimed at non-food items. The Direct catalogue and Extra stores supply anything from toys and electronics to furniture. The vast scope of Tesco's operations has meant that the retailer

[15] Sainsbury recently cut the value of its loyalty scheme by half during the spring of 2015. See for example an article by The Guardian, 'Sainsbury's Nectar points cut angers customers', 10th April 2015.

[16] Tesco offers its consumers the possibility to double the value of their vouchers and spend them on various different leisure activities namely, discounted meals, holidays and adventure parks among many others.

has compiled a rich set of consumer data. The Economist offers some insights into the data advantages of Clubcard to Tesco.[17] The Clubcard offers the retailer access to essential customer information and enables Tesco to channel personalized offers. For example, insurance companies are likely to rely on general demographic statistics to determine risk rates to set insurance product prices. On the other hand, the Clubcard allows Tesco to enjoy additional insights into consumer characteristics, consumption patterns and behaviour over time. Thus the retailer can discriminate between consumers and target low risk individuals with their insurance products.

Students at the London School of Economics, performed an experiment to demonstrate how Tesco uses Clubcard information when setting insurance premiums. The students firstly applied for car insurance with a blank Clubcard and then applied using their own personal Clubcards.[18] They received different insurance rate offers varying by as much as 18% in price, if compared to the benchmark data-free Clubcard. For example those who had never purchased alcohol using their Clubcard, received significantly lower quotes for car insurance. This simple experiment evidences Tesco's ability to leverage itself into other markets using customers' data obtained through a loyalty scheme. This aspect of loyalty schemes is beyond the scope of this paper and rather a direction for future research. Instead, we focus on the competitive aspects of the loyalty scheme itself and thus the next section considers the rationale for loyalty scheme strategies in the context of the lock-in effect and artificial switching costs.

1.3 Economic Theory of Loyalty Schemes

Throughout this section we present the main literature on switching costs and also look at the effects of loyalty scheme strategies in different types of markets. We focus on publications that look at competition in *nondurable* goods markets as this is consistent with the products sold to consumers in the groceries sector. In *durable* goods markets, for example for washing machines, retailers typically engage with customers on an infrequent basis. On the other hand, when selling perishable items such as fresh groceries, retailers engage in repeated interactions with consumers and therefore face different incentives to a durable goods seller. We also briefly touch on the effect of such strategies in intermediate markets to highlight the role of wider market features in

[17] 'Loyalty Rewards and Insurance: Every Little Helps', The Economist, November 5th 2011, (available online http://www.economist.com/node/21536605)

[18] 'Retailing: Spies in your Wallet', The Economist, November 5th 2011 (available online http://www.economist.com/node/21536604)

assessing the impact of loyalty discounts. The literature we present in this section indicates that when firms implement loyalty scheme strategies market outcomes and welfare effects depend on the number of factors. For example, the number of time periods entering the model, extent of product differentiation, number and symmetry of firms and ultimately, the design of the loyalty scheme itself (Caminal & Claici 2007, Fong & Liu 2011, Caminal 2012). On this basis loyalty reward schemes can either *soften* or *intensify* competition.

In presenting the literature in this section, we examine firms' strategies in the presence of switching costs more generally, and explain why firms have the profit incentive to create endogenous switching costs in some markets. We conclude by outlining a simple two period model à la Hotelling by Lal and Bell (2003) where grocery retailers compete through frequent shopper programs. This model does not explicitly model switching costs. Instead the authors rely on the Hotelling linear city framework to restrict behaviour of consumers depending on their position on the unit line. Due to the set-up of this model, consumers are assumed to differ in the way their behaviour changes in response to a repeat purchase discount. This is the exact type of assumption we propose to test empirically in the context of our empirical work. First, let us consider the fundamentals of switching cost theory.

Klemperer (1987) explains that consumers face significant costs of switching between brands in a variety of different markets. Switching costs which arise without the intervention of sellers are known as exogenous switching costs. The size of such switching costs arise *independently* to the firms' pricing and other strategic decisions (Caminal & Matutes 1990). For example, transaction costs are incurred by consumers when switching between providers *even* if both brands are entirely identical (Klemperer 1987). Typical examples include the transaction costs associated with switching bank accounts, mobile network providers or energy providers. Firms may also have the incentive to create switching costs through their own strategic behaviour (Klemperer 1995). For example, providers can affect switching behaviour of customers by entering into contracts (Fudenberg & Tirole 2000).

Alternatively, retailers can introduce strategies such as repeat purchase discounts and loyalty schemes to create artificial switching costs (Klemperer 1987). This type of strategy is the focus of this thesis and we note that loyalty scheme related switching costs are assumed to be *endogenous* as they arise as a result of firms' direct actions (Klemperer 1995). Thus, models of endogenous switching costs by definition assume that firms determine the size of switching costs themselves (Klemperer 1995, Farrell & Klemperer). Firms have the incentive to create these switching costs as they can lock-in consumers allowing the firm to enjoy market power over this segment of its customers (Klemperer1995). In this context, it has been argued that firms compete vigorously *ex ante* to gain and establish a large market share and are thus able to enjoy an *ex post* monopoly over locked-in consumers who face high costs of switching (Farrell & Klemperer 2007). We note this incentive applies to both exogenous and endogenous switching costs.

In light of the above, it can be shown that firms have a *profit* incentive to focus on growing and maintaining market share in the presence of switching costs. Further, considering multiple periods of competition, these incentives may also increase the intensity of competition between rival firms leading to lower average prices (Rhodes 2014). Thus, when considered in a multi-period dynamic setting, the effects of switching costs on consumers, firm strategies and equilibrium outcomes, depend on a number of factors. Below we consider a few scenarios which offer general insights on switching costs and associated firms' incentives. We then outline models of endogenous switching costs and of course those specific to loyalty schemes (point (v) below). Klemperer's (1995) provides a comprehensive literature review setting out the different ways switching costs can arise and their effects on market outcomes and welfare. Following Klemperer (1995), the main switching cost *types* are caused by the below factors:

i. Compatibility or interoperability of equipment create switching costs if the products are not interchangeable between different brands (e.g. pen and cartridge or computer hardware);
ii. Transaction costs lead to lower switching by consumers (e.g. cost of switching bank accounts or electricity provider);
iii. Costs of learning to use new brands (e.g. switching computer operating systems);
iv. Uncertainty about quality of untested brands create a perceived risk of switching (e.g. changing medicine);
v. **Discount coupons and loyalty rewards linked to repeat purchases**; and
vi. Psychological costs of switching related to noneconomic brand loyalty effects which alter consumers' preferences for known brands.

Firstly, let's consider the scenario where switching costs arise due to the learning involved in consuming an untested product (point iii above). This is somewhat similar, but distinct to search costs incurred when consumers search for a product to buy. Learning related switching costs arise where the consumer invests substantial time to learn how to use a product and faces switching costs with respect to switching to other products in the market (Klemperer 1987). This increases differentiation in the market and in this context, firms can achieve additional profit through greater market share. These types of switching costs arise even if *ex ante* the two products are otherwise identical (Klemperer 1987). In other words, in this example, switching costs increase differentiation due to psychological attachment consumers have with respect to well-known brands or products they have tried and tested in the past (Klemperer 1995). As switching costs increase perceived differentiation between brands and products, there is a strong relationship between a firm's current market share and its future profitability (Klemperer 1995). In fact, it can also be shown that markets with switching costs are more attractive to entrants than markets without consumer switching costs and can result in higher firm profits and prices (Beggs & Klemperer 1992).

With the knowledge that consumers are partially locked-in due to switching costs, firms tend to face the trade-off between *harvesting* consumers by charging higher prices or *investing* in market share through lower average prices (Klemperer 1995, Anderson & Kumar 2007, Rhodes 2014). In some cases this leads to a bargain-to-rip-off game where firms offer low prices to grow market share and then charge higher prices in future periods (Farrell & Klemperer 2007). In other words, anticipating that consumers will not switch away, a firm with a degree of market power can harvest customers who face switching costs by charging higher prices. Alternatively, it may seek to invest in growing its market share and therefore choose to reduce prices and avoid losing market share. This trade-off in the context of a dynamic model of multiple periods of competition can lead to ambiguous welfare outcomes. The outcome in terms of prices and welfare generally depends on the assumptions entering a model, namely the number of stages of competition (Rhodes 2014). Let us consider a few examples where exogenous switching costs lead to either *lower* or *higher* prices.

Villas-Boas (2006) considers a dynamic infinite period model of overlapping generations of consumers. Consumers and firms are forward looking, and these firms sell experience type goods. Thus in the model, consumers are uncertain about their future preferences and learn about a product only after consuming it. This scenario assumes switching costs of the type number (iii) above. Villas-Boas model shows that prices decrease to the extent that firms value the future. Thus in certain periods firms compete more aggressively and cut prices as this allows them to achieve higher market share in the future period. In the model, once a large enough number of consumers experience the good they are less likely to switch away. Then the firm exploits these customers and charges higher prices. In the model consumers only live for two periods. As a result, the price oscillates from low to high. The price is set lower when the firm seeks to incentivise consumers to try a product and then set high once a sufficient number of customers have experienced the product. If the model is extended to so that consumers live for more than two periods, higher prices may arise at equilibrium (Villas-Boas 2006). This is because the firm may have lower market share among "younger" customers, however, it will have a large market share of "older" customers who are locked in by learning related switching costs associated with experience goods.

Anderson and Kumar (2007) apply a two-period duopoly model where firms are asymmetric in their ability to attract loyal repeat buyers. This model is also extended to a model of multi-period competition and the results also hold under this extension. The scenario considered by the authors is based on empirical evidence of the relationship between well-known brands, greater customer loyalty and lower average prices. The assumption that a firm's pricing strategy increases loyalty of customers *endogenizes* the size of the firm's loyal base of customers. In doing so, the model creates a trade-off for the firms between *harvesting* old customers and *investing* in new ones. The authors show that as firms gain market share, they engage in more promotional activities and at equilibrium they offer lower average prices compared to rivals. In fact, Anderson and Kumar show that it is the "stronger" firm with the strongest brand and highest market share offers lowest average prices compared to the smaller rivals.

This result is contrary to what competing models of switching costs predict: firms raise prices as they gain market share due to the relative market power they enjoy over their customers who face either actual or perceived costs of switching (Klemperer 1995, Rhodes 2014). Rhodes (2014) also shows that switching costs can reduce prices paid by customers. The author applies a Hotelling model of infinite number of periods to show how switching costs redistribute overall welfare over time. His model assumes that firms are more patient than consumers which leads to lower prices. In this context, a firm's incentive to lock-in a given consumer is outweighed by that consumer's incentive to *not* be locked-in. Rhodes finds evidence of significant price heterogeneity in the short run in a dynamic setting. In addition, by assuming multiple generations of consumers, Rhodes is able to show that switching costs transfer welfare between different generations of consumers. Consistent with other models of switching costs, the firms find themselves in a prisoners' dilemma. In other words, it may be beneficial for them to make it more difficult for customers to switch due to the additional market share they can achieve. However, this behaviour also intensifies price competition and reduces the profits earned by competing firms (Rhodes 2014).

The above examples illustrate the many incentives and trade-offs faced by firms when consumers are locked-in through exogenous switching costs which arise *without* the intervention of firms. In particular we noted the effects on firm incentives associated with a "harvest" versus an "invest" strategy respectively. We note that a recent trend in the literature are models of switching costs incorporating a greater number of periods of competition (Rhodes 2014). This represents an important improvement because in reality, firms are generally forward looking and competition takes places over a long time horizon (Rhodes 2014). A multi-period model allows the researcher to better understand firm and consumer strategies in a dynamic setting. Preferences of consumers can change over time and firms may also change their strategies in response. Again, these types of effects can only be captured by considering multiple stages of competition. We note that this likely explains why more recent literature on switching costs incorporates multiple periods of competition revealing the diverse set of market outcomes in the presence of different types or a combination of switching costs. We now consider the main features of models of endogenous switching costs.

Point (i) above references the switching costs which arise as a result of incompatibility or interoperability of products. Setting the level of compatibility of products can also be an *ex ante* strategic decision by a firm and would therefore represent an endogenous switching cost. For example, tying pre-installed software to hardware can increase costs of switching for consumers to other types of software. Additionally, printers and ink cartridges or razors and razor blades are generally only compatible within the same brand. This incompatibility creates costs of switching for consumers. Knowing this, firms can make strategic decisions to make products compatible, or not. Matutes & Regibeau's (1992) relatively static duopoly model considers such a scenario and looks at the incentives of firms to standardize components. The authors assume that two firms produce differentiated products that are compatible components that can be combined into a system of products (e.g. computer and keyboard). Consumers obtain no utility from purchasing a single component, therefore firms have the incentive to produce compatible components.

The firm operating in a duopoly market sells either a system of their own components or offers components compatible with the rival firm. The Matutes and Regibeau paper considers these distinct scenarios and assumes that individuals value variety but have no brand preferences. Thus the degree of compatibility of components shifts demand in turn affecting the firms' strategies. The authors consider several pricing and product compatibility scenarios captured through bundling strategies and find that in most cases mixed bundling is at a least a weakly dominant strategy. By selling the components as part of a mixed bundle, consumer demand for the system of components increases due to the fact that individuals value variety. However, the firms would prefer not to offer a bundled discount as it would make them better off.

As pointed out by Klemperer (1995), in the scenario considered by Matutes and Regibeau, differentiation mitigates the anticompetitive effects of switching costs because due to their preferences for variety, consumers have the incentive to use more than one supplier. On the other hand, we also note that as in Anderson and Kumar's model with brand related switching costs, firms can rely on repeat purchase discount strategies to artificially increase the degree of perceived differentiation in the market and increase market power of firms (Klemperer 1995, Farrell & Klemperer 2007). In the case of endogenous switching costs, Farrell and Klemperer (2007, p. 2001) explain that there are also a number of other different incentives to consider namely that "[m]*arket participants may seek to either raise or to lower switching costs in order to reduce inefficiencies (including the switching cost itself), to enhance market power, to deter new entry, or to extract returns from a new entrant.*"

According to Klemperer (1995), "[t]*he simplest way to endogenize switching costs is to add to existing models* [which consider switching costs] *an initial ("zeroth") period, in which firms make compatibility or other choices that determine whether or not switching costs subsequently arise; we expect switching costs to be chosen where they raise future profits more than any current costs to firms of creating them.*" Thus the general framework assumes that the firm itself sets the size of switching costs *ex* *ante* and then competes against rivals. We now consider the seminal model of endogenous switching costs set out by Caminal and Matutes (1990).

The authors apply a two-period Hotelling model of endogenous switching costs and consider a differentiated product duopoly where firms can discriminate between new and repeat buyers. The authors compare outcomes under different price commitments and show the *type* of commitment in place matters for the market outcomes and prices paid by consumers. The authors find that under both coupons and price commitments, consumers pay decreasing prices. Price commitments themselves are shown to enhance competition and coupons, on the other and, tend to decrease competitiveness of markets. Caminal and Matutes explain that this outcome arises because a price commitment does not have an impact on the profits earned from loyal consumers in the second period and as a result, firms compete more aggressively in the second period. The implementation of a coupon which does not require a commitment on future prices is not costly to the firm as it can raise prices to compensate for the cost of the coupon. This reduces welfare of consumers paying higher prices.

We next consider an example of how a discount strategy creates an endogenous interdependence between demands for two unrelated products. In presenting this model we note that the same interdependence is created between time periods as a result of a repeat purchase discount. Gans and King (2006) evaluate the effects of a joint purchase discount for groceries and petrol in an oligopoly setting. The authors' model demonstrates how the strategy changes outcomes in the market compared to a uniform pricing strategy. The model assumes that the discount is determined *ex ante* (i.e. the model endogenizes the discount) and rivals respond in the next period. When rivals react, the discount softens price competition due to the prior commitment to offer bundled discounts. Overall industry profits are reduced if all firms resort to the same strategy representing another form of the prisoners' dilemma previously encountered. The results also indicate that the discount induces consumers to consume a brand mix that does not reflect their preferences. Thus the strategy is an effective tool used to increase loyalty of customers by increasing relative switching costs. The discount creates a strategic interdependence between otherwise independent purchases (Gans & King 2006).

Like in the above example of bundled discounts for petrol and groceries, multi-period models of competition allow the researcher to vary assumptions on the nature of price commitments in place. This applies to both models of exogenous and endogenous switching costs. It can be shown that price commitments may or may not be necessary in sustaining a competitive outcome in a dynamic setting (Farrell & Klemperer 2007). For example, there may be sufficient incentives in place for a forward looking firm to continue to compete vigorously against rivals without any price commitment being in place. In practice, firms can commit to future prices in different ways with varying effects as captured by the Caminal & Matutes (1990) model of endogenous switching costs. We next consider an extension of the above model which considers a dynamic model of competition in the presence of loyalty schemes.

Caminal & Claici (2007) model the competitive effects of linear and lump-sum discounts attributed to loyalty schemes and consider the effects on firms' market shares and social welfare. The model adopts a multi-period dynamic framework and is based on a groceries market characterised by monopolistic competition with free entry. The result shows that a loyalty scheme has pro-competitive effects if there are a large number of firms in the market and these firms are also able to commit to future prices. In the model firms initially face identical demand. However, when *ex ante* one firm decides it will offer a loyalty discount for repeat purchases this affects other firms' strategic decisions. In the model, a loyalty scheme strategy is the dominant strategy for all firms. As more consumers sign up to the scheme, the firm offering a loyalty scheme differentiates itself from rivals and increase its market share. This creates two competing effects, as discussed in the literature on exogenous switching costs.

On the one hand, firms compete more vigorously because they anticipate lower future equilibrium prices and fight for market share. We note that this set-up assumes that the firm is able to discriminate between groups of consumers. In other words, as consumers are locked-in by the strategy and without commitments to future prices, the firm can raise prices to past consumers while offering lower prices to newcomers. This result can be shown by extending the analysis to an overlapping generations model which shows that firms have the incentive to discriminate between new and old (repeat) buyers when they can differentiate between them. Thus with few competing firms the strategy can be shown to have anti-competitive effects. However, in most cases, Caminal and Claici argue that loyalty rewarding schemes produce pro-competitive effects so long as a commitment mechanism is in place. Thus, although the loyalty scheme is shown to influence demand and act as a business stealing device, the authors note that loyalty schemes typically increase social welfare and lead to lower average prices.

In Caminal and Claici's model it is assumed that firms can observe each other's pricing and future commitments to prices. This increases transparency in pricing behaviour in the market and increases the risk of collusion. Fong and Liu (2011) apply a dynamic overlapping generations model with an unlimited number of firms and show the conditions and loyalty scheme structure that tend to facilitate sustainable tacit collusion. The model specification implies that across time periods, firms can recognize their own repeat customers but do *not* differentiate between new customers and rivals' customers. The results the authors obtain build on the standard two-period model outcome. Compared to uniform pricing equilibria and two-period models of loyalty discounts, the authors demonstrate that loyalty rewards enable tacit collusion under both commitment and non-commitment to rewards and prices. Regardless of product and consumer heterogeneity, different loyalty rewarding pricing structures are shown to facilitate tacit collusion regardless of the market structure. The collusive outcome is sustainable for a wide range of discount factors when firms compete using different loyalty schemes. Here, without a commitment to future prices, the loyalty reward structure nonetheless results in a collusion sustaining discount factor and lower payoffs for deviating firms.

Basso et al. (2009) consider the moral hazard associated with loyalty programs and use a Hotelling duopoly model to show that the introduction of Frequent Flier Program ("FFPs") loyalty schemes can alter competition in the market. The authors find softening of price competition when firms implement the strategy. The authors show that FFPs soften competition rather than intensify it because prices and profits move in opposite directions. With a FFP in place, the airline can charge higher prices, while the more expensive the FFP is to operate, the higher the profits for the airline. This finding is also consistent with the fundamentals of switching cost theory outlined by Klemperer (1987, 1995). In other words, firms have the profit incentive to invest in creating endogenous switching costs in some markets as this allows the firms to enjoy greater market power over the segment of locked in consumers and to charge higher prices.

In the model firms choose prices and FFPs simultaneously and find themselves in a prisoner's dilemma. Both airlines hope that the other airline will choose not to operate an FFP in order to achieve maximum profit by being the only airline to offer an FFP in the market. Profits are lower if both airlines offer an FFP, and higher if neither of them do, even though equilibrium prices are lower. Although these equilibrium results are insightful, the authors' primary focus throughout the paper is to address how Frequent Flier Plans create a problem of moral hazard. The workers' demand becomes less elastic as their employer pays for the cost of the ticket, creating the moral hazard problem. The results show that the FFPs exacerbate this existing issue because they enable the airlines to charge higher prices by coercing business travellers to purchase more expensive tickets with side-payments via the FFP. The authors note also that their model "[…] *contrasts with the switching-cost approach in which the FFPs can, depending on the model, raise or lower prices and airline profits.* [They] *also showed that more costly FFPs may lead to higher profits than less costly plans, as they provide less efficient ways for firms to compete.*"[19]

Caminal's (2012) more recent paper addresses the alternative implication of loyalty programs, unrelated to endogenous switching costs. This approach considers the different design efficiencies of the reward schemes including first period lump-sum discounts and future price commitment designs. Caminal recognizes that private and social incentives may not align in real world markets where the "*[discount] policies are always less efficient than price commitment, and may imply even lower surplus than in the absence of behaviour based price discrimination* (Caminal 2012, p. 26)*"* This makes it particularly important to understand whether the type of loyalty program structure will reduce or improve efficiency. Although the model considers a monopolist firm, the reader is able to make inferences on the implications of the results under alternative market structures. For example, loyalty rewards can generate efficiency gains by encouraging consumer participation because "*consumers are willing to pay up-front for the promise of future low price* (Caminal 2012, p. 5).*"*

However, when considering a competitive setting across multiple time periods, Caminal (2012) argues that the design efficiency of a loyalty scheme will be difficult to measure because firms anticipate rivals' actions and multiple equilibria are likely to arise. Caminal (2012) assumes that the monopolist offers a 'contract' which bundles together the first and second period consumption, including the respective discount. Consumers are differentiated as being either first time or second time buyers. The firm is thus able to price discriminate between new and repeat buyers. If the monopolist is able to commit to future prices, there is an improvement in efficiency and total welfare

[19] Basso, L.J., Clements M.T., Ross, T.W., 'Moral Hazard and Customer Loyalty Programs', American Economic Journal: Microeconomics, Vol. 1, Nr. 1, 2009, p. 116

because consumer participation increases. More generally, Caminal argues that for the set of loyalty rewarding designs that encompass credible price commitments for future periods will improve the efficiency of the market equilibrium. However, if firms offer a discount without a price commitment in place, total welfare may not increase. While it may be difficult to quantify in practice, the loyalty reward scheme design has an important impact on the efficiency of the market equilibrium.

Many (albeit not all) of the models discussed thus far have applied some version of the Hotelling model. The Hotelling framework exposes consumers' relative preferences to shocks to account for differentiation and consumer heterogeneity. Caminal (2012) explains how "[*i*]*n this set up [loyalty rewards] allow firms to retain previous customers, even when rival firms offer goods or services that better match their current preferences. As a result, [loyalty rewards] are welfare reducing because they cause a mismatch in the allocation of consumers. However, in this view it is unclear whether [loyalty rewards] tend to relax or exacerbate price competition."* Therefore the insights offered by such models are subject to assumptions on consumer preferences and differentiation. Particularly as competition generally takes place along a far greater number of parameters than suggested by the Hotelling framework. This motivates us to focus our empirical work on the analysis of consumer preferences for loyalty schemes by considering the interaction of multiple dimensions of retailer characteristics.

Although our main focus is on loyalty schemes in the end-consumer market, there are valuable insights to be gained from analysis of loyalty discounts in upstream markets in terms of how the differences between these two environments drives outcomes. In upstream markets, multi-product suppliers can reward loyalty on both product combinations and quantities of goods purchased, much like in retail markets. In intermediate markets, firms can offer either *bundled loyalty discounts*, which are achieved by customers through purchases of multi-product bundles, or *loyalty rebates* that are defined as quantity based discounts (Greenlee et al. 2008). Considering the features of intermediate markets, the scope for exclusion as a result of contractual arrangements, which include discounts, can have a significant impact on prices in the long run if rivals exit the market. In upstream markets, the buyer-seller relationship is characterised by lumpy contracts and few interactions between buyers and sellers. Contractual arrangements between sellers and buyers can lead to artificially induced switching costs in consumers. Therefore in both upstream and end-markets, loyalty discounts can lock-in consumers.

In upstream markets, effects of loyalty discounts are associated with severe foreclosure effects akin to tying and bundled discounts and also pro-competitive effects in terms of lower prices and greater intensity of competition (Faella 2008, Elhauge 2009, Whish 2009, Economides 2010, Federico 2011, Zenger 2012). Loyalty rebates in general distort the competitive process by inducing loyalty via increased switching costs (Elhauge 2009, Faella 2008). Economides (2010) argues customers in these markets may perceive non-participation in a loyalty scheme as the equivalent to receiving a disloyalty penalty. Compared to the no-discount benchmark, loyalty rewarding schemes leave customer surplus unchanged because the firm benefits from locking in consumers and increases prices in future periods (Economides 2010). Thus the customers find themselves in a form of prisoners' dilemma. The customer could have been better off *not* participating in the loyalty scheme altogether if other customers signed up to the scheme. In addition, the commitment to future discount rates increases transparency which may facilitate collusion and lead to higher prices (Economides 2010).

The above models analyse firm strategies in the presence of one specific type of switching cost. However, it is possible that in some markets, consumers face endogenous and exogenous costs of switching between sellers or products (Shi 2012). Shi (2012) considers a two-period model of both exogenous and endogenous switching costs. The model is an extension of the above Caminal and Matutes (1990) model of endogenous switching costs. Shi (2012) extends the Hotelling model to consider the effects of exogenous and endogenous switching costs on market outcomes, which respectively, affect competition in very distinct ways. In the model, product differentiation is captured through transportation costs and firms set the size of endogenous switching costs through a loyalty scheme. Exogenous switching costs and the transportation costs entering the model have different effects on the size of endogenous switching costs chosen by the competing firms. As firms compete for market share through loyalty scheme strategies, endogenous switching costs are shown to increase in the presence of higher transportation costs. As part of this result, Shi finds that when transportation costs are high, consumers place more weight on the brand and are less responsive to loyalty schemes.

Thus there are two opposing effects to consider here. The brand effect arising from transportation costs on the one hand, and on the other hand, the cost and effectiveness of the loyalty scheme in attracting consumers. Shin shows that when exogenous switching costs increase due to the brand effect, loyal consumers are unaffected by a loyalty scheme. Due to the brand related exogenous switching costs, consumers do not switch and loyalty discount redemption rates are high. As such, the loyalty scheme becomes a costly consumer retention tool and on that basis, the level of endogenous switching costs set by the firm decreases at equilibrium. Shin demonstrates that both of these switching costs help retain customers and reduce brand switching in the market. In addition, Shi (2012) finds that a prisoner's dilemma arises at equilibrium because when both firms set higher endogenous switching costs they also lose more profits under this strategy. Shi's model underlines the complex nature of real world markets where consumers are likely to face opposing incentives when deciding whether or not to switch between retailers or products. We consider this important aspect of switching costs when looking at the competition policy framework for the assessment of markets with switching costs.

In light of the literature reviewed and presented above, the resultant effects of loyalty schemes on the market in question will typically depend on the market structure and the nature of consumer preferences. For example, it has been shown that in a duopoly, loyalty schemes are generally considered to be anticompetitive, while under monopolistic competition, loyalty schemes are associated with largely procompetitive effects (Caminal & Claici 2007). The role of consumer preferences is also important because the loyalty scheme has to increase switching costs for a large number of consumers to have an impact on competition through a lock-in effect. The lock-in effect itself then creates an artificial monopoly over consumers which can exclude rivals who are unable to compete with an equally attractive offer, either in terms of price, quality or other product characteristics.

On the basis of the above literature, we note a common assumption in models of both endogenous and exogenous switching costs. Keeping all else constant, these models assume that consumers incur switching costs in the same way. As far as we are aware, there exists only one recent publication which explicitly assumes that consumers have heterogeneous switching costs (Biglaiser et al. 2016). Biglaiser et al. (2016) consider a two-period model of exogenous switching costs under a duopoly. Unlike previous models, consumers are forward looking, have heterogeneous switching costs, make their choice of seller on the basis of price and the *type* of customer base the seller actually has. Thus, in this specific scenario, low switching cost consumers' behaviour is observed and followed by high switching cost consumers. The main result is that pricing decisions of firms are affected by the assumption of heterogeneous switching costs as well as profits and market shares. The authors provide an important contribution in that "[…] *heterogeneity of switching costs has complex strategic consequences which have largely been ignored in the literature. It will influence the strategies of firms, the equilibrium distribution of clients, and the value of incumbency*."[20]

Considering the above, we note the possibility that consumers may be heterogeneous in their switching costs and that *some* consumers may not perceive *any* switching cots when faced with repeat purchase discount. Lal and Bell (2003) apply this assumption to a variant of the Hotelling model used to analyse the impact of frequent shopper programs on market shares and profits in grocery retailing. We note however, that this specific model does not endogenize switching costs in the traditional way as described by Klemperer (1995). Instead, the model explicitly assumes that some shoppers are simply loyal and are unaffected by the presence of a loyalty discount. Oher shoppers are not loyal and are instead cherry-pickers who seek to find the lowest prices regardless of the brand.

Lal and Bell begin their paper by presenting the results of an empirical analysis of a product specific promotion on store profits. Using scanner level data the authors perform an empirical analysis to show that promotional discount strategies impact consumer behaviour and increase retailer profits. The empirical analysis shows that that high value customers who spend the most in store are the least impacted by the programs, so-called loyal customers. Also, the schemes have a positive effect on profit due to the impact on behaviour of those customers not classified as the loyal type. The "*empirical research suggests that supermarket frequent shopper programs, as currently implemented, are an attempt to get customers to spend more at a store in exchange for a discount—be it a ham, turkey or a discount.*"[21] We note that this overarching conclusion does not distinguish between loyalty schemes and promotional product discounts.

[20] Biglaiser, G., Crémer J., Dobos, G., 'Heterogeneous switching costs', *International Journal of Industrial Organization*, Vol. 47, p. 63, 2016.
[21] Lal & Bell (2003)

The authors also derive a theoretical model to explain the results of the empirical analysis. Lal and Bell note that their model is an extension from previous work in industrial organization on competition between two stores where consumers shop for a basket of goods.[22] The theoretical framework relies on the Hotelling model to represent consumer preferences along the usual linear city. The model is derived assuming a symmetric duopoly that incorporates consumer shopping costs. The authors provide an extension to include a loyalty scheme offered by a single firm and also where both firms offer a frequent shopper program. Below we present and explain the scenario derived by Lal and Bell (2003) where only one retailer offers a loyalty scheme as this best reflects the UK groceries market.[23]

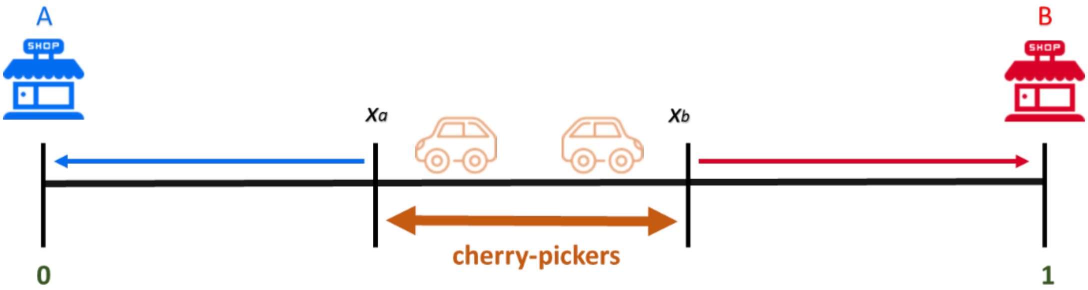
The analysis assumes that two supermarkets, *A* and *B*, are located at the two ends of a line of unit length and consumers are located uniformly along the line connecting the two stores. Consumers are distinguished by two specific behaviours; they are either loyal customers or cherry pickers. A loyal customer only purchases from retailer *A* or *B*, but never both, and cherry pickers look for the best prices and shop at both stores to achieve a saving on their basket price. This saving, $2d$, is achieved by cherry picking and purchasing the two baskets at two separate retailers rather than both baskets at one where a discount saving of d is achieved. Both retailers carry the same assortment of products as reflected in a typical basket of goods purchased by shoppers. The products available in store are assumed to be identical, however the prices of products are not always the same.

The two retailers, *A* and *B*, set prices of the items in the corresponding grocery baskets to signal a price image captured by P_a and P_b. The corresponding promotions d are assumed to be determined exogenously by the sellers of the products. The framework also assumes that consumers can obtain a repeat purchase discount L by purchasing both baskets of goods at retailer *A*. The model assumes that the grocery retailers incur zero marginal costs. Consumers incur a shopping/travel cost c per unit distance travelling to and fro from a store and their own location. This travel cost c captures the degree of differentiation, where the line connecting the two stores is a vector along dimensions differentiating the two stores. Figure 1.2 below is an illustration of the linear city framework used to analyse loyalty rewarding schemes in

[22] The authors refer to Lal and Matutes (1994) and Lal and Rao (1997).
[23] The UK groceries sector is characterised by the “Big Four” retailers, where two of the four offer a loyalty scheme.

Lal & Bell's (2003) model. The below diagram applies to the version of the model where a proportion of shoppers located between x_a and x_b are cherry pickers and shop at both *A* and *B* to obtain a discount 2*d*.

Figure 1.2 – *Hotelling linear city framework with cherry picking shoppers*



Firstly, we follow Lal and Bell and consider the result of the benchmark Hotelling model. Here the authors assume no loyalty schemes or price promotions i.e. *L*=0 and *d*=0. Assuming that the marginal consumer is located at distance *x* from store *A* gives the following constraint:

$$2P_a + 2cx = 2P_b + 2c(1 - x)$$

$$x = \frac{c + P_b - P_a}{2c}$$

Store *A* profits are:

$$2P_a x = P_a \left(\frac{c + P_b - P_a}{2c}\right)$$

Profits are maximised when:

$$P_a = 0.5(c + P_b)$$

The symmetric equilibrium implies the following prices and profits:

$$P_a^* = P_b^* = c$$

$$2xP_a = c$$

Considering the case of price promotions (where $d > 0$), retailers *A* and *B* sell two baskets each at full price P_a and P_b respectively, and two baskets at a discounted price $P_a - d$ and $P_b - d$. As noted above, the discount is determined exogenously by the manufacturers of the products. The model assumes that cherry pickers exist only if $x_a < x_b$, where the consumers located between x_a and x_b cherry pick between the two stores to and pay $P_a + P_b$ -2*d*. In order to ensure at least some cherry picking occurs, we must assume that the discounts compensate shoppers for the shopping costs they incur i.e. $d > c$. Also it follows that $d < 2c$ to ensure that not all consumers cherry pick.

If $d = 2c$, all consumers resort to cherry picking because the discount covers the shopping costs for all consumers, including those with the highest shopping cost of $2c$. Thus, in the model, cherry picking occurs only if $2c \geq d \geq c$.

We can define the relevant constraint that ensures cherry picking as follows. The consumer located at point 0 (where store *A* is also located) up to the consumer located at point x_a will be indifferent between shopping at store *A* only and cherry picking where: $[2P_a + 0 - d;\ P_a + P_b - 2d + 2\text{c}(1 - x_a)]$. The length of this and *A*'s resultant market share will therefore be as follows:

$$2P_a + 0 - d = P_a + P_b - 2d + 2\text{c}(1 - x_a)$$

Which gives:

$$x_a = \frac{P_b - P_a + 2c - \ d}{2c}$$

It follows that the area from the consumer located at point X_b where she is indifferent between cherry picking or buying only from *B* up to the consumer located at point 1 (where store *B* is located): $[2P_b + 0 - d;\ P_a + P_b - 2d + 2\text{c}x_b]$. The length of this distance and *B*'s resultant market share is:

$$2P_b + 0 - d = P_a + P_b - 2d + 2cx_b$$

Which gives:

$$x_b = \frac{P_b - P_a + c + d}{2c}$$

For the benchmark case $d = 0$, store profits were equal to c. In this case store *A* maximises:

$$\begin{aligned}(2P_a - d)x_a &+ (P_a - d)(x_b - x_a)\\ &= (2P_a - d)\left(\frac{P_b - P_a + 2c - \ d}{2c}\right) + (P_a - d)\left(\frac{d}{c} - 1\right)\end{aligned}$$

Differentiating the profit function with respect to P_a:

$$\frac{\partial\left[\left(\frac{P_b - P_a + 2c - d}{2c}\right) + (P_a - d)\left(\frac{d}{c} - 1\right)\right]}{\partial P_a}$$

$$= 2\left(\frac{P_b - P_a + 2c - d}{2c}\right) - \frac{1}{2c}(2P_a - d) + \left(\frac{d}{c}\right) = 1 + \frac{P_b}{c} - \frac{2P_a}{c} + \frac{d}{2c}$$

Setting the above function equal to zero and assuming the symmetric equilibrium dictates $P_a^* = P_b^*$ we obtain the optimal solution:

$$P_a^* = P_b^* = c + \frac{d}{2}$$

The optimal solution results in equilibrium profits for seller *A*:

$$\Pi_a = \left(2\left(c + \frac{d}{2}\right) - d\right)\left(\frac{2c - d}{2c}\right) + \left(c + \frac{d}{2} - d\right)\left(\frac{d}{c} - 1\right) = c - \frac{d}{2}\left(\frac{d}{c} - 1\right)$$

Compared to the benchmark result where $d = 0$, the profits here are less than c for $d > c$ and zero if $d = 2c$ when all consumers cherry pick.

Now we consider the scenario of price promotions and also that retailer *A* offers a loyalty scheme, i.e. $d > 0$ and $L > 0$. Therefore consumers who purchase both baskets at store *A* pay $2P_a - d - L$. Consumers located between x_a and x_b resort to cherry picking and pay $P_a + P_b - 2d$. In this scenario, for cherry picking to occur the discount must be greater than the travel cost and loyalty discount, $d > c + 0.5L$. However, travel costs must be sufficiently large to prevent everyone from cherry picking i.e. $c + 0.5L < d < 2c$. We can define the relevant constraint that ensures cherry picking as follows. The consumer located at point 0 (where store *A* is also located) up to the consumer located at point x_a will be indifferent between shopping at store *A* only and cherry picking which gives the following constraints: $[2P_a + 0 - d - L;\ P_a + P_b - 2d + 2\mathrm{c}(1 - x_a)]$. The length of this distance and *A*'s resultant market share will therefore be as follows:

$$2P_a - d - L = P_a + P_b - 2d + 2\mathrm{c}(1 - x_a)$$

Which gives:

$$x_a = \frac{P_b - P_a + 2c - d + L}{2c}$$

Similarly, consumers located closest to *B* purchase both baskets at store *B* while consumers between x_a and x_b cherry pick between the two stores. It follows that the area from the consumer located at point X_b where she is indifferent between cherry picking or buying only from *B* up to the consumer located at point 1 (where store *B* is located) i.e. $[2P_b + 0 - d;\ P_a + P_b - 2d + 2\mathrm{c}x_b]$. The length of this distance and *B*'s resultant market share is:

$$2P_b - d = P_a + P_b - 2d + 2\mathrm{c}x_b$$

The market share of *B* will therefore be:

$$x_b = \frac{P_b - P_a + d}{2c}$$

The firm then wishes to maximise the following:

$(2P_a - d - L)x_a + (P_a - d)(x_b - x_a)$

$$= (2P_a - d - L)\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) + (P_a - d)\left(\frac{2d - \ 2c - L}{2c}\right)$$

i.e.

$$\Pi_a = (2P_a - d - L)\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) + (P_a - d)\left(\frac{2d - \ 2c - L}{2c}\right)$$

Differentiating the above profit function with respect to P_a gives:

$$\frac{\partial\left[(2P_a - d - L)\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) + (P_a - d)\left(\frac{2d - \ 2c - L}{2c}\right)\right]}{\partial P_a} =$$

$$= 2\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) - \frac{1}{2c}(2P_a - d - L) + \left(\frac{2d - \ 2c - L}{2c}\right)$$

$$= \frac{-4P_a + 2L + 2P_b + d + 2c}{2c}$$

Differentiating the profit function with respect to *L*:

$$\frac{\partial\left[(2P_a - d - L)\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) + (P_a - d)\left(\frac{2d - \ 2c - L}{2c}\right)\right]}{\partial L} =$$

$$= -1\left(\frac{P_b - P_a + 2c - \ d + L}{2c}\right) + \frac{1}{2c}(2P_a - d - L) - (P_a - d)\left(\frac{1}{2c}\right)$$

$$= \frac{2P_a - P_b + d - 2L - 2c}{2c}$$

Setting the above first order conditions to zero and solving for P_a and L gives:

$$P_a = d + 0.5P_b$$

$$L = 1.5d - c$$

The optimal value of L implies that there is no more cherry picking as $2d - 2c - L < 0$ when $L = 1.5d - c$ and $d < 2c$. Here the profits to both stores are c which is greater than the previous case where the model was restricted to only price promotions. The authors explain that the scheme is most effective at changing the behaviour of cherry-pickers versus the behaviour of loyal customers. When one retailer offers the loyalty scheme, this enables the participating firm to gain consumers by compensating for individuals' 'shopping around' costs i.e. some consumers will no longer purchase from both firms and instead will purchase exclusively from one seller.

The behaviour of loyal customers on the other hand, who have a strong preference for either retailer *A* or *B*, is unaffected by the program. However these loyal shoppers still benefit from a loyalty discount by purchasing from a single retailer. In the model, the overall welfare change is due to a reduction in travel costs which is captured by the retailer in higher profits. However, this model does not account for differentiation between retailers, nor the dynamic effects of competition over time. Thus, Lal and Bell's model does not capture the potentially exclusionary effect of the loyalty scheme strategy. For example, in contrast, Caminal and Claici (2007) note that under a duopolistic market structure, loyalty schemes can create perceived switching costs for consumers which distorts competition in the market. Caminal and Claici explain that in duopoly setting, loyalty rewarding schemes are generally viewed as being anticompetitive.

Lal and Bell extend their model to address the setting where both firms offer a loyalty program. However, we do not provide a full derivation of the model extension as this variant assumes that *all* firms in the market offer a loyalty rewarding scheme. Instead, we provide a commentary on the key insights from the results obtained by the authors. When the model assumes that both retailers offer a loyalty scheme, cherry pickers are eliminated because their demand is entirely captured by one of the two retailers. Effectively, in this scenario the schemes cancel each other out, similarly to the outcome suggested by Caminal and Claici's (2007) model of monopolistic competition. Lal and Bell show that offering the scheme is no longer effective at enhancing profits compared to the situation where only one firm offered a scheme. When both firms offer the loyalty scheme, their profits are still equal to c.

The authors further extend the model to include two competing loyalty schemes and different customer segments. These consumers are differentiated by their travel costs c_1 and c_2. As before, "optimal" loyalty programs eliminate cherry picking because the reward for buying both baskets at the same store is set to be at or above *d*. Here the loyalty scheme itself plays a limited role in affecting behaviour. Store profits vary depending on the level of *d* and are maximised when one segment of consumers cherry picks between the two stores. The authors conclude that based on the results, it may be difficult to change the behaviour of already loyal customers. Therefore retailers need to either make their loyalty reward lucrative enough to sufficiently compensate consumers for shopping around costs, or alternatively, target specific customer segments whose behaviour can be materially influenced through a loyalty discount.

Comparable to its counterparts, the above model assumes a simplistic segmentation of preferences, even with the extension to include variable shopping costs c_1 and c_2. The model assumes that some consumers are loyal to a single retailer (maybe due to proximity or brand preference), others seek out the best promotional offers and the rest prefer a store offering a loyalty reward. From the above solutions, we can see that market shares and resultant profits are a function of several parameters including prices, promotions, shopping costs and the loyalty scheme. However, the model largely ignores the interaction between price and other dimensions of competition and differentiation in grocery retail markets such as product range, quality of service and quality of products. Therefore the extent to which loyalty schemes change consumer behaviour may not be fully captured by the model. It is far more likely that customers choose their preferred grocery retailer based on a wider combination of store characteristics in addition to the level of price.

We previously noted that this model is structurally different to the models of endogenous switching costs outlined further above. These models endogenize costs of switching in the zeroth period and the firms themselves set the size of the switching costs. On the other hand differences in preferences for repeat purchase discounts in the Lal and Bell (2003) model are captured through shoppers' relative positions in the Hotelling linear city framework. An interesting aspect of the above model is that it assumes that some consumers are not affected by the presence of a repeat purchase discount, while others are. In the Lal and Bell model, loyal customers remain loyal by definition *not* because they are locked-in through a repeat purchase discount. Perhaps they are locked in due to brand preferences but this point is not really considered by the authors. Instead, the model assumes that so-called cherry-pickers are impacted by the loyalty discount who shop around for good deals.

The Lal and Bell model is a highly simplistic view of the groceries sector which does not consider specifically the role of endogenous switching costs which arise due to the presence of loyalty schemes. However, the assumption applied does raise an important question. That is: do different *types* of consumers behave differently when there is a repeat purchase discount available? As stated above, to the best of our knowledge this point has not been explicitly addressed by models of endogenous switching costs in industrial organization. Instead, it is typically assumed that the firm can create artificial switching costs by offering repeat purchase discounts without considering the fact that this may not actually create switching costs for everyone. In other words, the literature to date largely ignores the *heterogeneous* nature of switching costs which may arise in some markets (Biglaiser et al. 2016). Further, the models unrealistically assume that all consumers redeem a loyalty discount if they visit the retailer offering a loyalty program. Instead, it is entirely possible that some consumers simply will not sign up to the scheme due to personal preferences and still visit the retailer offering the scheme on a frequent basis.

In light of the above literature, this thesis proposes the hypothesis that when firms implement loyalty schemes this will affect the behaviour of only a proportion of consumers in the population through artificially created switching costs. It follows that the behaviour of certain individuals will be *unaffected* as they prefer *not* to participate in any loyalty scheme at all, regardless of which retailer they choose. Thus their choice of retailer is independent of the availability of a repeat purchase discount. We discuss further below how empirical analysis can help inform and/ or test assumptions entering theoretical models. We next look at the empirical evidence on discounts tied to customer loyalty.

1.4 Empirical Evidence on Discount and Loyalty Scheme Strategies

This section presents empirical evidence on the effects of discount based pricing strategies adopted by firms, including loyalty rewarding schemes. We first consider the type of data available to retailers when making strategic pricing decisions and promotional strategies. We then present empirical evidence of how loyalty schemes based on repeat purchase discounts are likely to impact consumers' choice of seller and the prices paid by consumers. In doing so we note that few empirical papers assess the effects of loyalty schemes in dynamic retail markets. We note that this may be due to the difficulties in observing and quantifying switching costs. Farrell and Klemperer (2007, p. 1980) explain that because *"switching costs are usually both consumer-*

specific and not directly observable, and micro data on individual consumers' purchase histories are seldom available, less direct methods of assessing the level of switching costs are often needed." We also note that the majority of the existing empirical literature on endogenous switching costs focuses on frequent flier programs. In fact, the effects of frequent flier loyalty schemes are also driven by the well-known principal agent problem rather than just artificial switching costs faced by consumers. We keep this point in mind in evaluating the results of such models.

In marketing and operational research there are models for maximizing loyalty card and scanner data (Pauler & Dick 2006). For example, retailers can use the data to identify the bestselling items and target promotional activity accordingly. Further, the retailer can identify more profitable products that can be used to cross-subsidize discounted products (DeGraba 2006). In other words, this is a type of "loss leader" strategy where a retailer advertises one popular discounted product (the loss leader) but recoups the losses because consumers purchase other products during the same shopping trip (DeGraba 2006). More generally, loyalty cards and store scanner-level data offer a retailer revealed preference data on its consumers' shopping behaviour and respective sociodemographics. Considering the above, retailers like Tesco and others, can analyse extensive data to achieve optimal product offers, store-specific promotions and personalized discounts. Tailoring store offers to suit the most profitable consumer segments can maximize store profits (Pauler & Dick 2006). Loyalty schemes ensure a repeated interaction between the retailer and consumers to reveal essential knowledge of long term consumption patterns.

Consumers in the groceries market may face search costs as consumers may not be able to observe the quality of a product before consuming it, which is a feature of an experience good. Avery (1996) considers how this impacts the way consumers shop around for products, discounts and promotions. Avery (1996) performs an empirical analysis of survey data on the applicability of Stigler's theory of "Economics of Information". The focus of the paper is on the process of consumers' pre-shopping and in-store search activity in the US groceries market. In this scenario, before making a purchasing decision, consumers repeatedly engage in search activity to better inform themselves about products and their prices. This process improves purchasing outcomes for consumers, so long as the marginal benefit of search is at least equal to the cost of search. However, consumers differ in their preferences for search activity as a result of the underlying determinants of search costs. Following Stigler's definition, the magnitude of search costs depends on an individual's monetary situation, the opportunity cost of engaging in search and the transportation cost (Avery 1996).

The type of pre-shopping search activities considered in the Avery (1996) paper are coupon collecting, coupon swapping, tracking promotional activity and viewing various forms of advertising. The findings suggest that consumers engage in search activity to different degrees depending on their demographic characteristics. In addition, consumers are shown to be generally poorly informed about prices of products in store as suggested by previous research in the field. Avery argues that consumers are largely unaware of specific product prices and that instead, consumers focus on understanding the general pricing strategies and promotions of specific retailers (Avery 1996). In other words, the retail offers differentiated by retailer brand.

We also know that promotions are an important component of the competitive process in some retail markets. Volpe (2013) analyses promotion driven competition in an oligopoly setting by examining dynamics of pricing strategies of supermarket chains in the Unites States. Volpe's results show that the strategic promotional behaviour contributes to price variation in the groceries market in the US. Using data on prices and promotions from two major supermarkets, Volpe finds empirical evidence that the retailers will seek to match each other's promotional activity. The retailers seek to promote items that will incentivise consumers to switch from a rival store. This strategic firm behaviour is consistent with the evidence of greater more intense price competition in the presence of switching costs.

Consumers who are relatively price insensitive will likely not be influenced by a discount and will make their purchasing decision based on other dimensions of their preferences (Wang 2010). The retailer can focus on identifying patterns in product preferences among shoppers to personalize stores. Differences in consumer price sensitivity may not be perfectly observable to the firm. To allow for customers to self-select, firms can offer joint purchase discounts. Wang (2010) analyses data on bundled discounts for joint purchases of gasoline and groceries in Australia. He explores the retailers' motivations to offer the bundled discounts. In contrast to theoretic models, Wang did not find evidence for exclusionary conduct or predatory intent. Instead, he finds that the 'loss leader' advertising method is the most likely explanation for applying the bundled petrol discount to groceries. For petrol to be a profitable 'loss leader' advertising tool, there is a minimum spend attached to the rebate as we see in real life markets. Furthermore, Wang argues that consumer price sensitivity will affect behavioural decisions. Price insensitive consumers do not have the incentives to redeem their vouchers or join a loyalty scheme compared to price sensitive individuals. Thus price sensitivity, i.e. consumers' elasticities of demand, would need to be estimated to explicitly measure the effects of bundled discounts on profits.

Asplund et al. (2008) estimate cross-sectional data to identify presence of behaviour based pricing in the newspaper market in Sweden. The data shows that newspapers discriminate depending on the amount of competition they face in their local area. The authors argue that discounts targeted at rivals' customers increase the presence of price discrimination. For example, newspapers in locations with higher competition offer discounts to students. More generally, discounts are targeted depending on consumers' levels of price sensitivity. The results further evidence an existing relationship between market power and the value of the discount. Where the newspaper enjoys greater market power, a discount is not offered. In fact, the discount value is found to be inversely related to the firms' market share. Newspapers facing a higher number of rivals in their area offered a greater discount. While the data did not show evidence of switching costs, existence of such costs could be a potential explanation for the targeted discounts (Asplund et. al. 2008). These results suggest that discounts are associated with pro-competitive effects when there exist a sufficient number of competing firms in the market.

We previously discussed how some consumers join a loyalty schemes to avoid the so-called disloyalty penalty. There is evidence to suggest that this effect may also deter consumers from switching in future periods. Morell et al. (2009) present experimental evidence to support this claim. Morell et al. (2009) show that consumers who are subject to targeted discounts can make irrational decisions in future. The researchers perform a lottery style choice based experiment on a group of randomly selected individuals. Risk preferences and loss aversion statistics were calculated. The likelihood of switching to a different option was greatly reduced by participation in the rebate scheme. They explain this behaviour by the *Cumulative Prospect Theory* which predicts that targeted rebates harm consumers because they are less likely to be willing to switch to a better offer. Morell et al.'s analysis of targeted rebates, supports claims of competitive harm from targeted discounts creating perceived switching costs.

Let us also consider Hartmann and Viard (2008) who challenge the 'lock in' effect associated with shopping frequency reward programs. The authors argue that frequent shoppers do not experience high artificial switching-costs. They use data on

531 golfers, some of whom participate in a golf club loyalty program based on the 'buy 10 get 1 free'. The paper aims to measure switching costs by constructing a dynamic demand model with forward looking consumers. The approach relies on firstly deriving choice probabilities that incorporate the customers' expected utility based on the discounted value of different purchasing decisions. The specification is then refined to suit the specific loyalty scheme data which is estimated using random parameters logit (i.e. mixed logit model). The data and model allow measurement of elasticities of demand under the reward program and without. The loyalty reward scheme is shown to have no effect on the respective elasticity of demand of customers. Hartmann and Viard argue that customer's *ex-ante* valuations determine the effectiveness of the scheme.

This result is comparable to that of Lal and Bell's (2003) where loyalty schemes primarily influenced the behaviour of cherry-pickers. Hartmann and Viard suggest that frequent shoppers already have a brand preference for the product with the loyalty scheme attached. Because they are already assumed to be loyal customers, their behaviour is not influenced by the scheme. Based on the data, the impact on elasticities is akin to the firm offering equivalent price reductions absent the loyalty scheme. At the same time, we should recognize that golf enthusiasts will probably have stronger brand preferences for golfing courses compared with customers choosing between grocery stores. The authors offer an alternative explanation on the role of loyalty reward schemes unrelated switching costs. They suggest that their findings could be supported by loyalty schemes acting as a mechanism for volume related price discrimination that reduces uniform prices. Alternatively, they argue that loyalty reward programs could be a mechanism for exploiting the principal agent problem which has also been explored in the context of frequent flier programs.

An empirical research paper in marketing by Liu and Yang (2009) looks at how competing loyalty schemes influence individual program effectiveness. They focus on the US airline industry and perform a two-stage least square estimation with the value of sales as the dependent variable. Loyalty scheme effectiveness is shown to be determined by the relative market share of an individual airline. Thus suggesting that additional features associated with larger airlines, such as complementary resources, would enable them to obtain additional incremental sales from loyalty reward schemes (Liu & Yang 2009). It is widely acknowledged that airlines are subject to a costly minimum efficient scale due to significant economies of scale and network effects associated with operations. Thus the finding that airlines with high market share reap the greatest benefits from loyalty programs is consistent with airline market features (McCaughey & Behrens 2011). This is also consistent with the literature on switching costs which explains companies' focus on in investing in market share through strategies that include loyalty schemes and repeat purchase discounts. While previous studies suggested that competing loyalty schemes lead to a zero-sum game, the results do not support this argument (Liu & Yang 2009).While there is some evidence to suggest that market saturation will reduce loyalty scheme effectiveness, the effect is eliminated under high category expandability.

As a complementary analysis to the initial regression exercise, Liu and Yang (2009) estimate survey data using a two-stage panel regression analysis. The estimates provide a measure of attitudinal loyalty and the influence on loyalty scheme effectiveness, market share and scheme membership. The sample features 166 respondents' attitudes on 11 of the most recognized US airlines. Consumers are segmented by their preferences for category expandability. The authors define high category expandability as the airlines' ability to compete in other product markets. This feature improves the airline's competitive edge in the airline industry for consumers with preference for high category expandability. The results support the previous regression analysis that suggested that airlines with higher market share have a more effective FFPs. On the other hand, small market share airlines who offered loyalty schemes did not see their FFP have an influence over their members booking frequency. This effect was measured using a simulated scenario to compare members' and non-members' booking frequency for a particular airline.

In light of the evidence that larger airlines achieve greater gains from their FFPs, as compared to smaller airlines, antitrust concerns may arise if these airlines also enjoy hub dominance at airports. Prior to regulatory intervention, the first phase of analysis would have to seek to identify the causes of reduced competition at particular airports (Lederman 2007). FFPs could be seen as a mechanism to isolate participating airlines from intense competition at a given airport. The schemes alter behaviour and entice travellers to book flights which enable them to keep earing towards their FFP rewards (Lederman 2007). These effects must then be weighed against improvements in welfare arising from greater economies of scale and enhanced networks as many FFPs rely on partnerships between airlines and these agreements can have mixed welfare effects. As previously mentioned, achieving FFP scale in terms of additional airline partners, should improve the effectiveness of the program because of the importance of networks within the airline industry.

The distortionary effects on behaviour, can offer an airline further market power over segments of customers making FFPs a useful tool for airlines wishing to preserve their hub dominance. Using data on fares, passenger numbers and FFP scale over time, Lederman's (2007) paper looks at the relationship between demand, variation in an airline's dominance and FFP enhancements achieved through additional partnerships over time. The results present evidence that loyalty schemes can distort competition by influencing demand and equilibrium ticket prices. Unlike previous work, this empirical model enables Lederman to isolate the effects of changes to the FFPs and their resulting impact on demand at the airline's hub airport. The estimates show that FFPs can impact the equilibrium outcome both in terms of higher demand and fares. Enhancements to FFPs at airports where an airline is dominant, contributes to further increasing fares and passenger numbers.

Another study relies on actual FFP airline data and is carried out by McCaughey and Behrens (2011). The data is sourced from an anonymous US airline. The authors consider whether FFPs lead to members paying higher prices as a result of premiums. The results show evidence of behavioural effects associated with FFPs. The scheme's data shows that the airline is able to exploit different willingness-to-pay of travellers between the different tiers of the scheme (e.g. gold and silver membership) and to charge differentiated premiums. This result supports Lederman's (2007) estimates of increased equilibrium fares under FFPs and the principal agent problem of moral hazard explored by Basso et al. (2009). McCaughey and Bahrens argue that the optimal strategy for the airline would be to introduce even further tier segments within their FFP to fully exploit variation in WTP.

To analyse the airline FFP data, McCaughey and Bahrens apply discrete choice analysis using the mixed logit model. While there may be an additional computation burden, the approach allows to control for correlations between alternatives and individuals (McCaughey & Bahrens 2011). This among additional features, makes the mixed logit an attractive option for panel survey data as well as revealed preference data. The authors were able to estimate consumer behaviour attributed to different levels of program membership. The model also accommodated demand segmentation to identify preferences of specific demographic groups, namely income level, gender and FFP membership. Unlike other versions of the model, the mixed logit allowed

McCaughey and Bahrens to identify variation in taste over individuals as well as across different groups of individuals. Section 1.6 looks at the various applications of discrete choice models such as the mixed logit. Next, in light of the theoretical and empirical literature, we consider the competition policy considerations when investigating or assessing markets with the presence of either exogenous or endogenous switching costs.

1.5 Competition Policy & Switching Costs

The previous section presented literature on switching costs and loyalty rewarding schemes. Earlier models suggested that switching costs create poorer outcomes for consumers in terms of higher prices and lower welfare (Klemperer 1995, Rhodes 2014). We note however, that alternative models show how switching costs can intensify competition between firms and lead to lower average prices. This section considers competition policy in the presence of switching costs with a specific focus on loyalty scheme strategies. We begin by discussing if and when intervention is appropriate and in doing so refer to the recent Competition and Markets Authority ("CMA") investigations into energy and banking where switching costs were deemed to be particularly problematic. This section seeks to highlight the importance of adopting a consumer oriented competition policy in markets with switching costs and more generally. Additionally, we note that in reality, consumers are likely to face more than one *type* of switching cost when faced with choices between different brands (product or retailer). We therefore also discuss the effect that artificial switching costs may create on top of brand related switching costs, particularly in terms of locking in consumers and excluding rivals.

Even though it can be shown that switching costs intensify competition between firms in some situations, switching costs are generally assumed to be welfare reducing (Farrell & Klemperer 2007). Farrell and Klemperer (2007, pp. 2005-2006) argue that while switching costs may not strictly soften competition between firms competing across different factors, they are however, likely to make competition more fragile. Farrell and Klemperer note that on the basis of the evidence it is far more likely that switching costs are welfare reducing, result in markets performing less well and higher average prices. Thus they suggest that in some markets, intervention may be required to reduce switching costs. For example, in markets where firms strategically make products incompatible to increase their market power. In such situations, regulations may be required to discourage practices which seek to raise switching costs (Klemperer 1995, Gans & King 2001). We begin by considering switching costs which arise irrespective of firms' compatibility or pricing strategies. In such markets, carefully designed remedies to increase the rates of switching may be necessary to make markets more competitive (Waterson 2003).

As explained above, switching costs which exist in markets for various different reasons, allow firms to enjoy a monopoly over a segment of locked-in consumers. In turn, consumer behaviour and lack of switching makes firms compete less aggressively for consumers and reduces the intensity of competition in the market. However, competition policy historically focused on firms' behaviour and has sought to remedy abuse of dominance through behavioural or sometimes structural remedies (Waterson 2003). However, in some markets, intervention which targets consumers specifically may be required to improve the functioning of markets and outcomes for consumers (Waterson 2003). Waterson (2003) considers the example of the UK energy market *before* it was investigated by the CMA over ten years on.

On the surface, the market itself seems potentially competitive and there are a number of different suppliers. However, switching between providers is not common place and consumers are not necessarily informed about available alternatives, even if the information is publically available. In this case, Waterson notes that regulators should consider whether intervention may be required to improve switching in the market. For example through price comparison websites and through introduction of swift and low-cost switching services. We now look at the UK competition authority's intervention in two distinct markets with high exogenous switching costs, including the consumer retail energy market.

The CMA recently completed Market Investigations ("MIs") into energy[24] and retail banking[25]. We refer to these as the Energy MI and the Banking MI respectively. In the context of these investigations, the CMA considered the reasons for low switching rates in these respective markets and quantified average benefits of switching for different groups of consumers in terms of financial savings (Energy MI 2016, Appendix 9.2, Banking MI 2016, Appendix 6.2). The CMA concluded that low switching levels in both markets were one of the reasons resulting in overall poorer outcomes for consumers. In the context of its findings following these investigations,

---

[24] The final report and other materials relevant to the investigation can be accessed here: https://www.gov.uk/cma-cases/energy-market-investigation

[25] The final report and other materials relevant to the investigation can be accessed here: https://www.gov.uk/cma-cases/review-of-banking-for-small-and-medium-sized-businesses-smes-in-the-uk

the CMA proposed a number of remedies to encourage customer switching. This is consistent with the consumer focused policy considerations outlined by Waterson (2003).

In the context of these MIs, the CMA found a number of common features of consumer behaviour in these two markets. Let us consider these similarities. Firstly, in both markets the CMA identified low levels of switching. One of the explanations set out by the CMA, was that consumers were disengaged with retail energy markets (Energy MI 2016, para 125) and the CMA also found low levels of consumer engagement in the personal current accounts markets (Banking MI 2016, paras 64-66). This was found to be the case despite the benefits of switching identified and quantified in both of these markets. In the context of the Banking MI for example, some banks put forward the argument that low levels of switching was a reflection of high customer satisfaction. The CMA argued however, that given the gains associated with switching, in a well-functioning market with low switching costs, it would be expect to observe far greater switching by consumers of personal current accounts in the market (Banking MI 2016, para 6.25). In the context of the Energy MI, the CMA found that there were more severe issues in disengagement and customer response among prepayment customers compared to others. Differences in switching costs were also identified in the context of the retail banking investigation. In the Banking MI, overdraft users were found to be as likely to search as other consumers however, they were less likely to switch than other consumers (Banking MI, para 6.44). Both findings suggest that consumers can be heterogeneous in their switching costs for a variety of reasons.

In terms of the remedies proposed, the CMA noted that the detriment arising to consumers due to excessive prices in the retail energy market amounted to about £1.4 billion a year between 2012-2015 and that the detriment varied by customer group (Energy MI, paras 194-195). Thus among the proposed remedies, the CMA outlined a package of customer centric remedies to help "*customers engage to exploit the benefits of competition and to [protect] consumers who are less able to engage to exploit the benefits of competition.*"[26] This presents a shift in competition policy towards consumer focused remedies to improve the functioning of markets. In addition, the CMA proposed an entire current account switching package in the context of the Banking MI (paras 14.1-14.163). The package of remedies aims to improve the switching process,

[26] Energy MI 2016, para 206.

increase transparency between providers to help consumers inform themselves and raise awareness of the benefits of switching among consumers.

We note however, that the above market investigations looked at markets with very specific characteristics. Namely, high levels of market concentration and stable market shares over time. Further, in these markets providers sell experience goods where trust in the brand matters to consumers. In addition, proportionately large numbers of customers did not actively switch between providers. The CMA found that consumers were generally disengaged and typically uninformed about alternatives available. Thus there were a number of interrelated factors which exacerbated the negative effects associated with switching costs. However, absent some of the market features outlined above, switching costs may be less problematic. For example, in markets where the vast majority of consumers are active switchers, this may create the right incentives for firms to compete vigorously for market share. In this context, the decision of whether to intervene in a market where endogenous and/ or exogenous switching costs prevail, must be evaluated in the context of the wider aspects of a market, namely, price trajectory over time, distribution of market shares between firms over time and whether any one firm has significant market power. We now consider the treatment of endogenous switching costs from a competition policy perspective.

The two main competition concerns associated with loyalty schemes are interrelated. The strategy creates endogenous switching costs which can be shown to lock-in consumers thereby softening price competition and may also exclude rivals who are unable to compensate consumers through lower prices. The scope for exclusion in the context of repeat purchase discounts, arises because the consumer's desire to obtain a loyalty discount in future, creates an interdependence between purchase decisions over time. In turn, the seller establishes a sort of monopoly over its customers as they are locked in. This strategy may therefore exclude rivals in future time periods as consumers are unwilling to switch away. Consumers may find themselves in a prisoners' dilemma, whereby they would have been better off not participating in the scheme. We note that exclusionary conduct by a dominant firm can result in detriment to consumers' welfare through either higher prices, lower quality or lower innovation (EC Guidelines 2009, paragraph 19).

Prevention of exclusionary conduct by dominant undertakings is therefore a central element to competition policy. In reviewing the economic theory of loyalty schemes we noted both the exclusionary and pro-competitive aspects of the strategy.

In this context, the risk of exclusion or softening of price competition, should be balanced against any pro-competitive effects that may arise due to the strategy, in particular where it intensifies competition between rival firms. For example, Gans and King's (2001) consider a model of different regulatory regimes in the context of technological endogenous switching costs. Gans and King (2001) argue that regulators face a trade-off between imposing regulations which reduce switching costs against the *amelioration costs* which arise as a result of the intervention. The authors' model shows that ameliorating switching costs results in lower prices, however, consumers or firms may end up absorbing these costs, mitigating the positive effects of the intervention. Thus we note that any intervention into markets of endogenous switching costs, needs to be balanced against the likely benefits of the remedy and who is likely to absorb the costs of the intervention.

We now look at the likely competition implications of endogenous switching costs where other switching costs are already present in the market, say those related to brand effects. As found by Shi (2012), the presence of exogenous switching costs may reduce incentives for firms to create endogenous switching costs. This occurs because the effect of exogenous switching costs outweighs the effect of endogenous switching costs. Consumers redeem their loyalty discount and the firm makes lower profits as a result. We note the possibility that this effect could go both ways. For example, if we were to assume a scenario where endogenous switching costs outweigh the effects of exogenous switching costs, firms' would have the incentive to offer customers repeat purchase discounts. This is the case of the Caminal and Claici (2007) model where in the context of homogenous products, each forward looking firms' dominant strategy is to offer a repeat purchase discount. Depending on the number of firms in the market and the type of price commitment, the strategy results in either a procompetitive outcome (large number of firms) or an anticompetitive outcome (small number of firms).

On the basis of the discussion above, we note the importance of considering the wider features of a market in the assessment of firm strategies. In this context, different types of empirical analyses can supplement the insights offered by theoretical models, albeit these approaches also face certain limitations (Waterson 2014). For example, competition authorities can empirically assess the effects of business strategies through a counterfactual. However, establishing an accurate counterfactual can prove challenging or impossible due to the complex nature of real-world markets and the plethora of economic variables to consider (Economides 2010, Greenlee et al. 2008). More generally, empirical techniques in competition cases have been steadily evolving and are becoming an essential component to an effects based approach. In this context, market definition and merger control assessments have become increasingly more reliant on sophisticated empirical techniques (Lianos & Genakos 2012). These techniques are also being applied outside of merger control to different competition policy areas (Lianos & Genakos 2012).

It is also widely recognized that different sectors have distinct features and important nuances. In this context, empirical methods allow the research to understand such differences (Waterson 2014). For example, the same strategy may have very different effects depending on the type of market under investigation. In this context, empirical modelling would reveal such idiosyncrasies. We note also that the researcher can turn to discrete choice experiments ("DCEs") to inform an assessment of a very specific competition question. For example, DCEs can be applied to test an assumption entering theoretical models involving consumer choice (Farrell & Klemperer 2007). Following Farrell and Klemperer (2007, p. 1980) [w]*here micro data on individual consumers' purchases are available, a discrete choice approach can be used to explore the determinants of a consumer's probability of purchasing from a particular firm*." We consider this approach further in Section 1.6 below where we also introduce the concept of discrete choice experiments, including setting out some examples of its applications. We also explain the nature of trade-offs involved when adopting such an approach compared to theoretical modelling. In doing so we also highlight the additional insights which can be achieved through a discrete choice experiment, in particular in the context of informing the assumptions entering a theoretical model.

1.6 Applications of Discrete Choice Analysis

Drivers of consumer decision-making have long interested researchers across disciplines (Louviere et al. 2000, pp. 20-21, Keane & Wasi 2013). Discrete choice models can be used to analyse consumer behaviour using actual consumer data or experimental data collected using some form of instrument. For example, one of the empirical papers we presented above, applied the mixed logit model to analyse the effect of frequent flier programs participation on choice of air fares using actual airline passenger data (McCaughey & Behrens 2011). This model is one of the more flexible alternatives available to researchers. We note however, that in practice it may be difficult to obtain micro-level consumer data. In such cases, the researcher can design an experiment for the collection of micro-level consumer data which can also be estimated using a discrete choice model (Louviere et al. 2000, pp. 20-21).

We note that one of the applications of empirical techniques in the context of switching costs, is to test aspects of theoretical models (Farrell & Klemperer 2007). This specific application is the focus of this thesis and below we outline a few examples of discrete choice models to show how discrete choice methods can be applied to the assessment of different types of hypotheses on consumer choice. The examples of discrete choice modelling we present below are intended to demonstrate the flexibility afforded to the researcher to test different hypotheses on consumer choice. The underpinnings of discrete choice models can be traced to Lancaster (1966) who suggested that individuals gain utility from the characteristics and features of goods.

Tirole (1988, p. 99-100) explains how the Lancastrian approach differs from traditional models in industrial organisation that assume a form of vertical and/or horizontal differentiation. Instead, Lancaster's framework assumes that "[g]*oods are defined as bundles of characteristics, and the consumers have preferences over characteristics. The consumers may have heterogeneous preferences over characteristics.*" DCEs typically require carefully constructed instruments through which data can be collected on the population of interest. Data can be collected either through a lab experiment or a survey for example. This data can then be fitted to a number of different discrete choice models allowing for the estimation of preferences. For example, the mixed logit model allows the researcher to test for presence of preference heterogeneity between individuals.

We note that applications of choice modelling stretch across disciplines, including but not limited to, marketing, transportation studies, migration economics, environmental economics, and health economics. Let us consider some examples of applications of discrete choice experiments to different contexts. In marketing research, discrete choice modelling assists in the optimisation of advertising strategies, allows demand forecasting for new products and has numerous other applications (Keane & Wasi 2013). Seetharman et al. (2005) present a review of applications of multi-category choice models used in marketing research. A number of the models outlined in the paper rely on discrete choice modelling. Using scanner-level (revealed preference) data, these models help explain different shopping outcomes by estimating consumer purchasing patterns across different categories of products.

Chung and Rao (2003) adapt the nested logit model to measure preferences for bundle attributes. The authors implement a survey and collect a sample of 136 undergraduate students. The data is then used to measure preferences for a computer hardware bundle. This particular model informs the researchers about the optimal design for a pure bundle of a system of complementary goods. Each bundle is defined by a set of observable attributes which vary across alternatives. This framework allows estimation of consumer willingness-to-pay for the bundle attributes. By applying a latent class specification as part of the estimation procedures, the authors are able to identify heterogeneity in consumer segments. The model is a type of generalized extreme value ("GEV") model also in the logit family of models.

Le Cadre et al. (2009) rely on a variant of the nested logit and estimate consumer preferences for French telecom operator bundle offers. They note that most research assumes that the consumer valuations are known to the firm, while in the real world, consumer preferences may not be clearly observable. Motivated by the gap in literature, the authors construct an approach to model consumer preferences for service bundle offers. They use data from an extensive questionnaire of 1014 families in France. Part of the questionnaire asks consumers to answer demographic questions and to grade bundle offers in terms of their reservation prices and attribute ratings. This 'grading' approach is the basis of latent class modelling that allows for segmentation of data based on taste heterogeneity. The findings allow the authors to define an optimally priced bundle and identify more profitable market segments.

Using questionnaire data from 1000 Seoul households, Shin et al. (2009) construct a GEV nested logit model and estimate consumer heterogeneity and preferences for telecommunications services. The research concentrates on bundling from a consumer welfare perspective rather than a marketers'. Their model allows them to adjust for different preference assumptions from economic models of commodity bundling. The empirical work presents estimates of willingness-to-pay and welfare gains for consumers when choosing the bundled goods. This approach also allows Shin et al. to isolate the effect of a bundled discount on consumer behaviour.

The work outline above mainly relies on the standard logit or variants of the nested logit specification. The main downside of standard logistic regression models, is that they require for the *Independence from Irrelevant Alternatives* ("IIA") condition to hold. This is because an individual's choice of alternative will most likely be correlated over an unobserved factor i.e. variation in taste. By implementing the mixed logit, the researcher can help overcome this limitation. In addition, this specification allows the researcher to estimate heterogeneity in preferences, segment the data based on observable characteristics and accommodate various substitution patterns in decision making (Hensher & Greene 2003, Train 2009). For example, using the mixed logit, Hess and Polak (2005) model passenger preferences for airport choice in regions with multiple airports. They refer to a sample of passenger survey data on 5097 respondents which they segment by traveller type (e.g. business vs. leisure). Previous models on airport choice applied the standard CL, leaving a gap in the literature on variation in taste within and across market segments (Hess & Polak 2005). Hess and Polak improve on previous approaches by implementing both the standard logit and mixed logit models to capture preference heterogeneity. Their research paper offers important guidance on the pros and cons associated with applying the mixed logit.

Wine economists, Bonaria Lai et al. (2008), implement both the mixed logit and multinomial logit to estimate taste heterogeneity for Sardinian wine. The data is collected using an online survey with a consumer sample of 138 wine drinkers. The authors champion the mixed logit model for its flexibility in measurement of preferences and choice. On the other hand, Dahlberg et al. (2012) perform an analysis of local migration data on 1444 individuals in Sweden to estimate preferences for public services. The approach uses the Stata *mixlogit* command by Hole (2007). The estimation enables the researchers to identify preferences for community characteristics and to test for heterogeneity in these preferences within and across different demographic groups.

Using a sample of 557 Israeli households, Blass et al. (2010) apply an 'elicited choice probability approach' to measure preferences for electricity reliability. The authors derive a linear version of the mixed logit model which requires a minimal computational effort compared to the non-linear version. Even with the random parameter specification, this particular approach does not necessitate simulation methods. The authors outline a novel, less cumbersome, alternative compared to stated preference approaches (Blass et al. 2010). The survey respondents evaluated a series of hypothetical electricity bills differentiated by their respective characteristics. However Blass et al. (2010) required survey respondents to perform an additional task. Unlike the usual stated preference approach where respondents choose one of several alternatives, Blass et al. ask respondents to note down percentage grades to indicate the percentage likelihood of choosing either alternative. This approach is in contrast to a standard practice which requires respondents to only have an option to state 0% or 100% per alternative i.e. either the chosen or rejected alternative (Blass et al. 2010). The analysis enabled the researchers to identify differences in taste, derive individuals' preferences and WTP of individual respondent segments.

In the field of health economics, Regier et al. (2009) apply the mixed logit model and compare several estimation techniques. The authors use survey data on individuals' choices of genetic testing alternatives. Since data estimated with simulation methods may not always converge to a maximum, Regier et al. apply both the maximum simulated log-likelihood and Hierarchical Bayes ("HB") procedures. The HB procedure allows the authors to verify that the classical approach of maximum simulated likelihood converges to a global maximum. We note that the mixed logit also accommodates willingness to pay estimation which has important uses in healthcare economics. In this context, Hole (2008) applies the mixed logit to measure patients' preferences and willingness-to-pay for general practitioner appointment characteristics. Hole (2008) uses data of 409 respondents in the UK, collected using a survey constructed with a *D*-optimality algorithm.[27] The data is fitted to a number of model specifications including the standard multinomial logit, latent class and mixed logit. Because demographic variables enter the estimation process as interactions with the primary explanatory variables, Hole also tests different versions of these models, with and without interaction terms.

The examples we presented above show how DCEs enable the researcher to address specific consumer oriented policy and behaviour questions. In addition, in some markets consumers are central to firms' strategies and the competitive behaviour (Waterson 2003). For example, it can be shown that low levels of switching in consumers can reduce the competitiveness of markets and lead to poorer outcomes for consumers. However, theoretical models in industrial organisation typically focus on the behaviour of firms, ignoring important aspects of *consumer* behaviour (Waterson 2003). Thus, competition policy and the economic models used to study competition issues, should closely consider consumer behaviour in such markets (Waterson 2003). In this context, discrete choice experiments offer the tools to assess the assumptions entering theoretical models. Namely the drivers of consumer choice and determinants of switching and equally non-switching.

[27] Chapter II of this PhD thesis contains an appraisal of different techniques in survey design.

A discrete choice experiment typically presents respondents with a survey containing a series of questions that mimic, to a certain degree, a multi-period model of competition. A typical assumption in a DCE is that consumer preferences are constant at the level of the individual, thus the data only represents a snapshot in time. While a DCE approach is unable to capture changes to consumer tastes over time, it can identify differences in tastes between groups of consumers. It should also be noted that in the context of various types of experiments, it is challenging to recreate complex real world markets and to engage participants who reflect typical consumers (Waterson 2014). In addition, consumers may not always respond in the same way as they would in real world markets. For example, consumers may state they like high quality, when in fact price is by far the most important determinant of choice. Care must therefore be taken in designing a discrete choice experiment used to estimate preferences in the population. We dedicate the next chapter to this area and in doing so we outline our preferred discrete choice model, the mixed logit, and consider different approaches for the collection of data.

1.7 Conclusion

The literature review presented in this chapter focused on the role of switching costs in different types of markets and theoretical model set-ups. In Section 1.2, we outlined features of the UK groceries market where a number of retailers offer customers loyalty schemes. This type of market represents an environment where consumers face artificial costs of switching due to the strategic behaviour of competing retailers. Section 1.3 considered the role of exogenous and/ or endogenous switching costs in determining firms' strategies and outcomes in a retail market. On the basis of the literature presented in this chapter, we found that costs of switching may either soften or intensify competition. However this generally depends on a number of factors, like the number of firms competing, the number of periods of competition entering a model and the nature of price commitments in place.

We found that more often than not, switching costs are associated with welfare reducing effects. When consumers are locked-in, due to actual or perceived switching costs, firms have a profit incentive to invest in market share and exploit such exogenous switching costs. We also found that in some markets, firms have the profit incentive to create endogenous switching costs through strategies incorporating repeat purchase discounts and loyalty schemes. Such models assume that firms decide the size of endogenous switching costs *ex ante* then compete against rival firms. Further, we found that heterogeneity in switching costs has largely been ignored in the theoretical literature. This is the case even if switching costs (exogenous or endogenous) themselves can been shown to have a direct impact on firms' strategies, market shares and anticipated profits. More generally we found that the effects of loyalty schemes are not well known in the literature, particularly from the perspective of the consumer as theoretical models focus on the strategies of firms.

Sections 1.4 and 1.6 presented some examples of applications of discrete choice methods, including experiments, for modelling determinants of consumer choice, highlighting the versatility of the approach across different markets. In light of the evidence presented in this chapter, including the competition policy considerations set out in Section 1.5, we propose a discrete choice experiment to model preferences of consumers. This type of approach is adaptable to different settings and can assist the researcher in testing the assumptions which enter theoretical models. In this spirit, we propose to test empirically whether consumers differ in the way they incur artificial switching costs due to their heterogeneous preferences for loyalty scheme strategies. In doing so we are able to determine whether consumers are likely to be heterogeneous in their costs of switching when retailers offer repeat purchase discounts. The empirical evidence on consumer preferences in the UK groceries market presented in the third chapter of this thesis, suggests that consumers differ in their costs of switching when firms implement loyalty schemes, with at least some consumers' choice of retailer being completely unaffected. Our findings have direct implications for the theoretical models used to analyse these strategies which have generally ignored this aspect of the market.

Chapter II

Methodology for a Discrete Choice Experiment

2.1 Introduction

This chapter outlines the methodology for a discrete choice experiment on consumer preferences for loyalty scheme discounts and other grocery retailer features. The experiment is inspired by theoretical literature on endogenous and exogenous switching costs which prevail in a number of different markets. We performed a review of the literature in the first chapter and found that firms have the incentive to create artificial switching costs in some markets by introducing loyalty scheme strategies. However, the models used to analyse loyalty schemes typically assume that consumers will react in the same way to the strategy. In other words, these models assume that consumers are homogenous in how they incur artificially created switching costs. We argue that this is unlikely to be the case in real world markets because consumers have diverse tastes more generally. We propose instead that consumers are heterogeneous in their switching costs when firms implement loyalty schemes.

In reality, some consumers simply will not care about receiving a repeat purchase discount while other consumers on the other hand, may choose a retailer specifically on the basis that they offer a loyalty scheme. We therefore propose to challenge the assumption that consumers are homogenous in switching costs artificially created by retailers. This chapter outlines a *D*-efficient survey designed on the basis of actual grocery retailer features in the UK groceries market. We estimate the data collected as part of this process using the flexible mixed logit model. In doing so we overcome some of the limitations of theoretical models discussed in the previous chapter. The approach in survey design outlined in this chapter focuses on obtaining robust parameter estimates and we therefore outline a survey designed following efficiency design theory. The results of the empirical analysis are presented in the next chapter.

When revealed preference ("RP") data is unsuitable or unavailable, researchers, marketers and regulators can follow a stated choice approach to analyse the effects of different business strategies on consumer choice and demand. DCEs rely on a stated preference ("SP") approach for data collection and the analyst can choose between different discrete choice models to estimate the data once it has been collected (Louviere et al. 2000, pp. 20-21). By undertaking a stated choice ("SC") approach, the researcher can re-create true market scenarios and produce the necessary data to model consumer preferences, estimate substitution patterns between alternatives and undertake forecasting procedures (Louviere et al. 2000, pp. 51-65). Therefore a wellexecuted discrete choice experiment ("DCE") can inform the researcher on a range of issues. For example, in the UK, the competition authority uses surveys to collect data recognizing that:

> "*…when conducting competition investigations and the evidence from these surveys is an important component of its findings. Survey evidence is also proving to be useful in remedies work.*" (Competition Commission 2010)

Rather than assuming a specific type of consumer, DCEs allow the research to measure various types of behaviour, both rational and irrational. Stated choice studies have been a popular tool amongst researchers because of their ability to generate data that captures realistic market decisions which can be analysed using a discrete choice model (Huber & Zwerina 1996, Louviere et al. 2000, p.1). The process of implementing a stated choice experiment requires carefully constructed hypothetical choice scenarios to present to study participants (Bliemer & Rose 2009). The generated questionnaire typically presents respondents with 2-4 hypothetical scenarios (i.e. choice situations) and these questions are typically presented in survey format.

In responding to the survey, participants choose between *alternatives* which are distinguished by their features, otherwise known as *attributes*. The research much choose the number of alternatives and corresponding attributes to enter the design (Bliemer et al. 2008). Attributes of available choice alternatives will differ in their dimensions, referred to as *levels of alternatives*, such as different prices or levels of quality. Respondents must choose their preferred alternative from the options presented to them based on these observable characteristics and trade-offs which define each available option. Typically respondents are faced with trade-offs between higher prices and higher quality versus lower prices and lower quality for example. The sample size requirement, number of choice situations and the unique combinations of attributes and their levels to create a questionnaire, are typically drawn from an underlying *experimental design* (Bliemer & Rose 2009).

The quality of the experimental design itself therefore drives the precision and statistical significance of parameter estimates when performing the empirical analysis. In other words, there exists a relationship between the statistical properties of stated choice experiments and the econometric models used to estimate the experimental data. Designs which have been derived on the basis of the statistical properties of discrete choice models are called "efficient designs". Here, efficiency refers to the minimisation of the standard errors of parameter estimates. Researchers have been increasingly relying on efficiency based designs, namely *D*-efficient designs (Huber & Zwerina 1996, Bliemer et al. 2008). *D*-efficient designs are increasingly used in applied research because their purpose is to minimise the standard errors of the parameters at design stages and improve the quality of the results obtained when estimating parameter values. This is also our chosen approach in the context of the pilot and final survey designs.

In this chapter we present the evidence showing that *D*-efficient designs offer empirical advantages over traditional *orthogonal* designs. For example, we explain that Daniel McFadden's (1974) conditional logit, and the extension to the flexible mixed logit model can be accommodated at design stages through a *D*-efficient design. We also note that by introducing the empirical model at design stages we are able to achieve improvements in the precision of parameter estimates when fitting the data to different specifications (Bliemer & Rose 2009). In presenting the benefits of efficient designs, we also outline certain trade-offs compared to orthogonal designs, in particular that the researcher must assume prior values for the mean coefficients of the variables of interest.

This chapter is structured as follows. Section 2.2 discusses the role of SC studies in applied research. In doing so we highlight the versatility of the approach compared with other methods. Section 2.3 focuses on the econometric models that we propose to use to estimate our data. We derive McFadden's (1974) conditional logit model and the extension to our preferred model, the flexible mixed logit model. We explain how both these models can accommodate stated preference data and allow the researcher to measure behaviour using maximum likelihood estimation. After reviewing our preferred econometric models, Section 2.4 outlines the trade-offs between efficiency based designs versus traditional orthogonal designs.

We then outline additional considerations in survey design in Section 2.5 and Section 2.6 considers different methods available to draw a representative sample of the population of interest. We also outline the biases associated with different types of survey data collection methods. Throughout Section 2.7, we present the qualitative evidence on the groceries market we collected to determine the relevant attributes and their respective levels to enter the experimental design. In doing so we present the design of the pilot survey also in Section 2.7 and discuss the results in Section 2.8. In Section 2.9, we outline the final survey design, explain the underlying considerations in the design of survey questions on respondents' sociodemographic characteristics.

2.2 An Introduction to Stated Choice Methods: Theory & Practice

Over the last 50 years, various institutions have to an extent relied on stated choice data where revealed preference data was unavailable, inaccessible or simply unusable. Stated choice methods provide the tools to assess the effectiveness of business strategies, evaluate policy decisions and to forecast demand for new products (Louviere et al. 2000, p. 21). In this context, a stated choice experiment can be the optimal method to obtain panel-level data and measure consumer behaviour and preferences. This approach allows the researcher to isolate the independent influence of variables on some observable outcome (Bliemer & Rose 2009). The process typically involves asking a sample of respondents to answer a sequence of questions in a survey format. These questions typically ask the study participants to choose between alternatives that are differentiated by specific distinguishing features that the analyst has chosen to enter into the underlying experimental design.

The approach allows for the collection of a rich and unique data set to model individuals' preferences for different product or service features. For example, choice experiments allow forecasting procedures, can provide information on willingness-to-pay ("WTP") estimates for new service improvements and can assist in the formulation of policy design. In this context, stated preference methods have been applied across a variety of sectors such as environmental economics, healthcare, marketing, and transportation studies to name a few. The UK's competition authority has also frequently implemented SC studies during market investigations and merger reviews (Competition Commission 2010, pp. 44-62)

Nonetheless, some economists argue that revealed preference data is strictly superior because it reveals what people *actually do* instead of what they *say they would do* (Louviere et al. 2000, p. 21). However, both revealed and stated preference data are subject to their own advantages and limitations and must thus be used in context of the empirical investigation at hand. Even if revealed preference data is available, it may not be particularly useful or appropriate for the researcher (Louviere et al. 2000, pp. 20-24). In real world markets, explanatory variables of interest such as price, may not produce sufficient variability over time to allow for an estimation of preferences. In addition, variables of importance can be highly collinear in real world markets which creates problems for data estimation. In light of these statistical considerations, even if we had access to revealed preference data on grocery retailer choice, properties like collinearity and autocorrelations in the explanatory variables could hinder our ability to accurately isolate the determinants of behaviour in the market. Carefully constructed stated preference experiments founded on economic principles of choice behaviour will produce data that is equivalent to RP data (Louviere et al. 2000, p. 21). The data obtained from a stated choice experiment can overcome statistical problems associated with RP data and will be equally suitable for the same econometric models that measure discrete choices for RP data.

In light of the advantages and versatility of stated choice methods for recreating true market scenarios, we argue that economists should most certainly be interested in stated preference techniques (Louviere et al. 2000, p. 21). Economic principles of consumer choice theory are rooted within stated preference methods. These principles are founded on the paradigm of choice that is integral to choice experiments.[28] This includes both the design of the choice experiment and the various choice models that can be used to estimate the SP data (Louviere et al. 2000, pp. 2-3). The consumer's choice paradigm found in DCEs was first proposed by Lancaster (1966). Lancaster (1966) offered a novel approach to measuring the utility gained from consuming different goods and services. Prior to Lancaster's contribution, goods were assumed to be the objects of utility itself. Instead, Lancaster (1966) proposed the novel idea that individuals gain utility from the characteristics and features of services and/or products. More precisely the elements of the paradigm of choice define:

> "*…the function that relates the probability of an outcome to the utility associated with each alternative, and the function that relates the utility of each alternative to a set of attributes that, together with suitable utility parameters, determine the level of utility of each alternative.*" (Louviere et al. 2000, p 34)

Incorporating the paradigm of choice to a stated choice study requires that the available alternatives are differentiated by a set of characteristics that provide varying levels of utility to the consumer. The recorded sequence of choices made by different respondents provides information about the relative importance of alternatives and attributes i.e. characteristics. Following the choice paradigm described above, the researcher can estimate the utility associated with different alternatives and their respective characteristics by deriving the probabilities of choice. The alternatives chosen by a given study participant are assumed to provide them with the highest level

[28] Please refer to Louviere et al. (2000) Chapter 1 for a complete overview of the paradigm of choice and its relationship with discrete choice models.

of utility compared to the rejected options. Due to the nature of choice experiments, stated preference studies will produce discrete choice data of either ranked choices or alternatively, chosen or rejected options i.e. 0/1 outcome. Preferences can then be estimated using discrete choice models which are founded on the principles of random utility theory.

Random utility theory states that behaviour in the population can be defined by a broad set of individual behaviour rules. This relationship can be represented by an indirect utility function that is derived below in this section. This equation contains a random component that represents the unobservable determinants of choice such as random taste variation in the population (Louviere et al. 2000 p. 34). The observable components of choice are the attributes and levels of chosen alternatives. The general random utility model that captures choice behaviour can be adapted to derive the family of discrete choice models. When formulating the underlying experimental design to populate a survey, the researcher must specify the indirect utility function which contains the attributes and levels that define available alternatives. Subject to the study objectives, the collected data can be estimated by selecting the preferred functional form from a range of discrete choice models like the conditional logit and mixed logit models (Louviere et al. 2000, p 34). The different functional forms found in existing discrete choice models, produce their own respective choice probabilities that accommodate different behavioural specifications. Hence, the choice of the econometric model that relates utility to estimated choice probabilities is integral to the experimental design of the study. Consequently, the researcher must select a choice model based on the type of data that will be generated by the SC study.

As mentioned above, the experimental design refers to the matrix of values that are used to generate the final survey questions (Bliemer & Rose 2009). A given design matrix contains the different combinations of attributes and their respective levels, as specified by the researcher in the indirect utility function. Attributes enter the model as the explanatory variables and the different combinations of attributes define the differences between alternatives that are presented to respondents. The researcher can specify the dimensions of the attributes and decide the number of alternatives that the respondent will face in any given choice task. For example, price and quality attributes vary in magnitudes and can be used to describe the features of a given alternative. Each alternative in a given survey question will be distinguished by the specific combinations of levels of prices and quality i.e. their magnitudes. The combinations of attributes and levels are contained within a design matrix which is then used to populate the survey. The columns of the design matrix contain the specific alternatives and their respective attributes that are varied within the design. The rows on the other hand, each represent a distinct survey question.

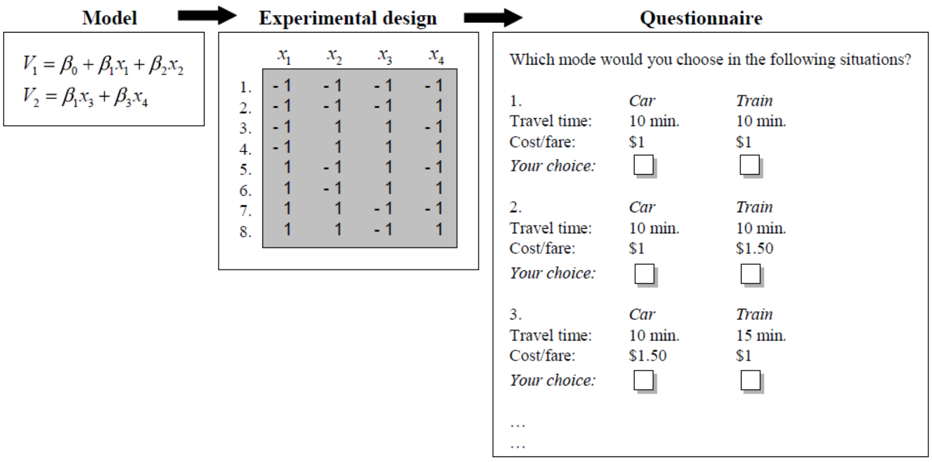
The general process of creating a survey for a discrete choice experiment is outlined in Figure 2.1 below. "*Independent of how the matrix is set out, the experimental design performs the same function, that being the allocation of attribute levels to choice tasks*..."[29] The design in Figure 2.1 is only an example of a possible matrix structure, which can take one of two forms. The first approach, which is also the example matrix below, assumes that each row represents a different choice situation and each column is a different attribute within the experiment. In this case, groups of columns form different alternatives within each choice task. The other possible approach assumes that each row of the matrix is an individual alternative while each column represents a different attribute. For this type of matrix format, multiple rows are combined to form a single choice situation.

The example presented below follows the first approach (rows indicate choice situations) and assumes a design with 2 attributes, travel time and cost/fare that take on two levels each. First, the researcher must specify the indirect utility function for each alternative, "car" and "train" i.e. specifies functions V_1 and V_2. The attributes can then be matched with a specific alternative. This is achieved by adjusting the utility function of the model presented in Figure 2.1 below. The β coefficients in this particular example capture the mean effect on utility of alternative-specific attributes x_1, x_2, x_3, x_4. In the example below, x_1 represents the travel time associated the alternative "car", while x_3 is the travel time associated with choosing "train". Then, x_2 represents the cost of driving a car and x_4 indicates the fare associated with travelling by train. The researcher must also specify if the attributes are *generic* or *alternative-specific*. In this example, β_1 is generic as it appears in both utility specifications, while the other βs are alternative-specific. β_1 captures the mean effect of travel time on utility, while β_2 represents the effect of car travel cost on utility. Following the collection of survey data, the model parameters β that capture the mean coefficients for "car" and "train" can be estimated by the researcher using a statistical software package.

[29] Bliemer, M.C.J., Rose, J.M., 'Constructing Efficient Stated Choice Experimental Designs', *Transport Reviews*, Vol. 29, No. 5, 2009, p. 588

Once the utility functions have been specified as described above, the design matrix is populated using a specific type of coding scheme. "*The most common ones are design coding (0, 1, 2, 3, etc.), orthogonal coding ({-1,1} for two levels, {-1,0,1} for three levels, {-3,-1,1,3} for four levels, etc.), or coding according to the actual attribute level values.*"[30] In the below example the experimental design coding is orthogonal. In this example, each attribute has 2 levels respectively. These -1/1 values are later replaced with actual attribute levels of the design. Looking at the first row of the experimental design, a "-1" for x_2 indicates a cost of $1 for the car alternative. The third row contains a "1" under x_2 to indicate a cost level of $1.50 for choosing a car. For a "good" design, the levels of attributes must vary sufficiently between alternatives and across the choice tasks to isolate the effect on utility. Effectively, the research wants to know the contribution that each attribute has on the overall level of utility. The underlying experimental design therefore determines the quality of the survey data and precision of parameter estimates. The final data set is the product of any design assumptions, such as the attributes and attribute levels chosen by the researcher.

Figure 2.1 – *Experimental design process*



(Ngene Manual 1.1.1, ChoiceMetrics 2012, p. 57)

The stated preferences methodology outlined by Louviere et al. (2000, p. 255), states that the researcher must firstly define the unique study objectives of the choice experiment. The study objectives must be supported by a combination of relevant theoretical, qualitative and quantitative evidence. This facilitates appropriate selection of the key attributes that are known to have at least some effect on consumer behaviour and the choices they would make in the real-world. After defining the study objectives,

[30] Ngene Manual, version 1.1.1, 2012, p. 59

researchers can choose how attributes are allocated within the design matrix by following one of several methods in experimental design theory. Typically, researchers choose to adopt either an orthogonal or efficient design approach. The designs primarily differ in the assumptions imposed by the analyst, specifically on the type of correlation structure between attributes in the design matrix. We discuss this in greater detail further below.

Let us consider the trends in the literature in the context of the *types* of experimental designs being used for survey data collection.

A common application of SC studies is to design policy and/or evaluate the efficacy of policies in achieving company objectives. This requires an understanding of the expected consumer valuation for products that are not yet available on the market. In light of increasing global pressures to reduce carbon emissions, transportation economists have turned to discrete choice experiments to analyse preferences and prospective demand for alternative fuel vehicles (Achtnicht et al. 2012, Hackbarth & Madlener 2013)**.** A recent study by Hackbarth and Madlener (2013) is one of many published papers that contributes to the already vast DCE literature in transportation studies. The sample of 711 German drivers answered a web based survey on choice of vehicle. The questionnaire presented the participants with 15 hypothetical choice scenarios in addition to sociodemographic questions. The authors' choice experiment applies a mixed logit specification which allows them to explore expected willingness-to-pay for different vehicle features and simulate how changes in these characteristics will likely affect market shares. The findings from the study seek to improve policy design by the German government to effectively shift households' consumption towards more fuel efficient automobiles. However, the authors do not provide further information on the chosen experimental design for the study.

Choice experiments are also prevalent in healthcare studies because they are effective at addressing various policy issues (Bekker-Grob et al. 2012). DCEs in healthcare were first used to value utility enhancing features of patients' experiences, such as waiting times or friendliness of the staff. Applications of SC studies to date have stretched across a wide range of policy issues. For example, using an orthogonal design, Marti (2012) performs a discrete choice experiment to determine preferences for smoking cessation treatments. The 131 selected respondents who answered the questionnaire were cigarette smokers and were also asked to provide sociodemographic information. The respondent information was used to segment preferences using the mixed logit model. The orthogonal design produced a total of 16 choice situations that were divided into two separate blocks of 8 questions. The two blocks were then divided among the respondents. Therefore, an individual respondent faced 8 hypothetical choice scenarios on their preferred medications, which were differentiated by measures of price, expected side-effects and drug effectiveness.

Bekker-Grob et al. (2012) undertook an extensive literature review of DCEs in health economics. The authors find that the primary aim of choice experiments in healthcare economics is to derive monetary measures as performed by Marti (2012). Such measures include WTP estimates for different medical products and services to measure the relative importance of time, risk and health outcomes as captured by the underlying differences in features of the alternatives presented to respondents. The studies reviewed by Bekker-Grob et al. (2012) that cover relevant information on experimental design, have predominantly used variants of orthogonal designs. More recently however, there has been a shift towards the implementation of DCEs that use a *D*-efficient design strategy. The authors' review of the literature demonstrates that health economists are continuing to contribute to the ongoing evolution in experimental design theory by progressively acknowledging the importance of introducing design efficiency in the construction of their DCEs.

Instead of relying on traditional orthogonal designs, researchers are making increasing use of more flexible econometric models and state-of-the-art *D*-efficient designs. Louviere et al. (2011) and Bekker-Grob et al. (2012) stress that the lack of detailed publications on DCE methodology and lack of best-practice guidelines remains one of the biggest challenges for applied researchers undertaking a stated choice study. This issue arises because there is no "one-size-fits-all" approach in the design of discrete choice experiments. This creates confusion over the optimal choice of design for a given context and publications frequently omit information on the experimental design used to collect the data. This is problematic because the experimental design plays a pivotal role in the accuracy and effectiveness of the stated choice study.

In addition, best-practice guidelines are ever evolving, thus making them a moving target for researchers (Louviere et al 2010). This is largely due to the fast paced evolution of this dynamic field, where inevitably state-of-practice lags behind the approaches that are currently state-of-the-art (Louviere et. al. 2010, Bekker-Grob et al. 2012). As a result, there is limited guidance on the appropriateness of outlining specific behavioural and statistical assumptions in a given choice experiment (Louviere et al.

2011). Throughout different fields of study, academic literature on DCEs rarely offer sufficiently detailed information on why specific assumptions were imposed, nor do they reveal important details on the overall process taken for design generation. To mitigate these risks and to improve the accuracy of the results at the estimation stage, the researcher must have well-defined research objectives and gather appropriate qualitative evidence at survey design stage (Bekker-Grob et al. 2012). In addition, there are a number of dedicated researchers whose work focuses on methodological concerns in choice modelling experiments (Louviere et al. 2000, Bliemer & Rose 2009, Scarpa & Rose 2008, ChoiceMetrics 2012). In addition, we refer to the UK's Competition Commission's (2010) guidance report addressing methodological concerns in this area. Combining the sources of evidence outlined above we note the following essential steps in determining the survey's experimental design (unrelated to sampling and data collection methods). The researcher must:

(i) define the study objectives;

(ii) select the econometric model that will be used once data is collected;

(iii) choose between creating a labelled or unlabelled SC experiment;[31]

(iv) perform qualitative research and undertake a pilot study to select DCE alternatives, attributes and levels to include in the design matrix and define the number of choice situations to present to study participants;

(v) evaluate, compare and select an experimental design that incorporates all the desirable properties and required assumptions to achieve the study objectives outlined during the first stage of the experiment.

(Louvière et al. 2000, Bliemer & Rose 2009, Scarpa & Rose 2008, Competition Commission 2010)

In the first chapter we addressed point (i) by defining our study objective: evaluate, in a realistic setting, whether loyalty scheme discounts determine (or not) consumers' choice of retailer in the UK groceries market. The next section addresses point (ii) above. We discuss alternative discrete choice models and in doing so derive McFadden's (1974) conditional logit and provide an extension to the mixed logit model by following Train (2009). We conclude that the mixed logit ("ML") is our preferred

[31] A labelled choice experiment defines alternatives with a name or brand. A labelled alternative would be defined as "German car" or "American car". Alternatively, an unlabelled choice experiment defines alternatives as A, B, C…etc. or 1, 2, 3…etc.

specification for the estimation of data because it is the least behaviourally restrictive compared to the alternatives.

2.3 Selection of the Econometric Model

When choosing between econometric models at survey design stages, the researcher must decide which model will be best suited to achieve the research objectives (Train 2009, p.19). The different models used in DCEs are differentiated by their distinctive choice probabilities that are used to estimate the data (Rose et al. 2008). Hence different types of data will be more suited for a particular discrete choice model. We note that the methodology outlined in this paper is formulated for both the mixed logit and McFadden's (1974) conditional logit models as they share a common functional form.

The conditional logit ("CL") model is the so-called workhorse of discrete choice models as it has been by far the most widely used in DCEs (Louviere et al. 2000, p. 65). In the literature, researchers sometimes refer to the conditional logit model as the multinomial logit model because functionally, the multinomial logit model can be expressed as the conditional logit. While the conditional logit model has been widely used, it suffers from certain restrictive properties such as the key assumption that consumers are homogenous in their preferences. On the other hand, the mixed logit can be adjusted to approximate any choice model by varying some of the assumptions imposed (Train 2009, p. 19). Hence, the main distinction between these two models is that the mixed logit offers significant advantages in terms of flexibility in accommodating a variety of preferences. Advances in computational capabilities paved the way for ever increasing use of simulation based methods in discrete choice modelling as required by the mixed logit model (Train 2009, p.134). Below we explain how the conditional logit can be extended to the mixed logit specification i.e. random parameters logit, to accommodate less restrictive statistical properties.

Following McFadden (1974) and Train (2009), we derive the conditional logit and the extension to the flexible mixed logit. McFadden's conditional logit model can be derived by assuming Lancaster's objective characteristics interpretation of utility. Recall that this approach assumes that the characteristics of products and services generate utility for consumers, not the product or service itself. When using survey data, each individual choice situation faced by the respondent is treated as a choice moment at time t where $t = 1,\ldots, t$. Data constructed from a stated choice experiment represents a form of panel data because each hypothetical choice scenario is treated as a moment in time. The choice probabilities are then derived by capturing that variability in preferences over the repeated set of finite choices. We provide the extension of the model to panel data after firstly deriving the basic behavioural model below as defined by Train (2009).

Both the conditional logit and mixed logit models are part of a wider family of random utility models ("RUMs") and can be derived following the same approach. RUMs are founded on the principles of ordinal preferences with origins from the Neo-Classical theory of individual choice (Bately 2008).[32] McFadden adapted the RUM to practical applications and "*reconstituted RUM from a model of an individual engaged in repeated choices, to one of the choices of a population of individuals.*"[33] The intuition behind the approach can be described as follows. Firstly, we assume that the researcher observes that decision maker n chooses between available alternatives J. However, we also assume the researcher is unable to observe the actual amount of utility the decision maker gains from making that particular choice. Instead, the researcher observes the set of choices made by individual respondents over the set of choice situations. By selecting a specific alternative, the individual obtains a level of utility that can be expressed by the observable characteristics of that chosen alternative. This utility can be defined as U_{nj}, where $j = 1,\ldots,J$. Therefore, the choices made by respondents represent relative utility differences between alternatives and the different attributes, instead of absolute utility.

The underlying approach assumes utility maximising behaviour, whereby the respective decision maker will choose the alternative that maximises his or her utility. The decision maker knows the utility he or she obtains, however this is not observable to the researcher. The behavioural model is derived by assuming that alternative i will be chosen if and only if, $U_{ni} > U_{nj} \; \forall \, j \neq i$. The set of observable attributes faced by the decision maker and also observed by the researcher, can be represented by a vector $x_{nj} \; \forall \, j$. In addition, the researcher can observe some of the decision makers' individual characteristics denoted as s_n. Thus we can derive a functional relationship which relates the observable factors to the decision maker, with a function: $V_{njt} = V(x_{nj}, s_n)$. This function is known as the representative utility function.

[32] Please refer to Bately (2008) for further information on the theoretical underpinnings of RUMs.
[33] Bately, R., "On Ordinal Utility, Cardinal Utility and Random Utility", Theory and Decision, Vol. 64, Issue 1, 2008, p.1

As pointed out by Train (2010, p. 15), V will often depend on parameters that are unknown to the researcher and must be estimated using quantitative methods. As the researcher cannot observe utility in its entirety, it is assumed that $U_{nj} \neq V_{nj}$. However, utility can be decomposed to derive a decomposed utility function: $U_{nj} = V_{nj} + \varepsilon_{nj}$, where ε_{nj} represents the unobserved segment of utility that is part of U_{nj} but not captured by V_{nj}. We treat ε_{nj} as random because it is unknown $\forall\, j$ alternatives. The random component represents the unobserved influences that are also determinants of choice. The joint density of random vector $\varepsilon'_n = (\varepsilon_{\mathrm{n}1}, \dots, \varepsilon_{nj})$ is a function $f(\varepsilon_n)$. This density is necessary to make probabilistic statements about decision makers' choices. The density function can be derived by assuming that the probability that the decision maker *n* chooses alternative *i* is*:*

$$
\begin{aligned}
P_{ni} &= \mathrm{Prob}\big(U_{ni} > U_{nj}\ \forall\ \mathrm{j} \neq \mathrm{i}\big) \\
&= \mathrm{Prob}\big(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj}\ \forall\ \mathrm{j} \neq \mathrm{i}\big) \\
&= \mathrm{Prob}\big(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}\ \forall\ \mathrm{j} \neq \mathrm{i}\big)
\end{aligned}
$$

Using the density function of the unobserved portion of utility we can derive the probability density distribution expressed as an integral:

$$
P_{ni} = \mathrm{Prob}\big(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}\ \forall\ \mathrm{j} \neq \mathrm{i}\big)
$$

$$
\int_{\varepsilon} \big(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}\ \forall\ \mathrm{j} \neq \mathrm{i}\big) f(\varepsilon_n) d\varepsilon_n
$$

Here $I(\cdot)$ is the indicator function which equals 1 when the above expression holds true and 0 otherwise (Train 2009, p. 15). Hence the survey response data must also be arranged with "0" representing the rejected options, while "1" indicates the chosen alternatives. Different assumptions about the distribution of the unobserved portion of utility determine the resulting model and functional form of the underlying choice probabilities. Following Train (2009, pp. 34-37), the random term in the logit model assumes a distributional property derived in McFadden's (1974) seminal work. The random portion of utility $f(\varepsilon_n)$ follows an identically and independently distributed ("iid") extreme value distribution, or otherwise known as the Gumbel and type I extreme value distribution. The distribution is a limiting distribution for an increasing sample size that describes the distributions of maximum and minimum values of a sample of independent and identically distributed random variables. The solution to the above integral is the formula for the logit model given by a closed form expression (Train 2009, p.36):

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_J e^{V^{nj}}}$$

Following Train (2009, pp. 50-51) the expression can be extended to panel data, of the type that can be generated using surveys. Assuming individual n chooses alternative *j* in a given choice moment (or survey question) *t* we have the utility function $U_{njt} = V_{njt} + \varepsilon_{njt}$ and the choice probability can be expressed as follows:

$$P_{nit} = \frac{e^{V_{nit}}}{\sum_J e^{V^{njt}}}$$

Recall that we are not measuring the absolute levels of utility here; the value of importance is the relative utility being measured using the unique set of choices made by individual respondents. To measure utility, the researcher must establish what is actually observable when estimating the data. A stated choice experiment requires specification of observable characteristics i.e. attributes, their levels and final choices of alternatives to present to respondents. Recall that the researcher observes the choices made by respondents and the alternatives chosen by respondents are defined by the corresponding attributes of alternatives found in vector x_{nit} which corresponds to individual *n*. Thus the observed portion of utility is $V_{nit}(\beta) = \beta_n' x_{nit}$, where β_n is a vector of individual-specific coefficients, or so-called weights, that are part of the observed portion of utility $V_{nit}(\beta)$. The choice probabilities can be solved using maximum likelihood because they are globally concave in parameters β (Train 2009, p. 37). With the assumption that $V_{nit}(\beta) = \beta_n' x_{nit}$ we can reformulate the above logit probability:

$$P_{nit} = \frac{e^{\beta_n' x_{nit}}}{\sum_J e^{\beta_n' x_{njt}}}$$

To summarize, when applying the standard logit formula, the researcher estimates the utility individual *n* obtains from choosing alternative *j* at choice situation *t* given by:

$$U_{njt} = \beta'_n x_{njt} + \varepsilon_{njt}.$$

The observed portion of utility $V_{njt}(\beta) = \beta'_n x_{njt}$ represents the utility achieved given the chosen alternatives *j* by an individual respondent *n* at choice moment *t*. The expression allows estimation of parameters of utility captured in β based on the chosen alternatives presented in a questionnaire, and these will be captured in the vector of attributes x_{njt} for each individual. The primary advantage of the standard logit formula above is that it can be solved analytically. However, the conditional logit suffers from certain restrictive properties. For example, the choice probabilities imply proportional substitution between alternatives and assumes the property of independence from irrelevant alternatives ("IIA") (Train 2009, pp. 42-46). The IIA condition states that an individual's choices will only depend on available alternatives in terms of the *relative* odds of choosing either alternative. The logit probability imposes a constraint that does not allow for variation in consumer's choice in the face of additional alternatives and their respective attributes. Most importantly, the standard logit will not be able to accommodate random taste variation because it assumes that preferences are homogenous across the population. Train (2009) argues that researchers aiming to capture additional variation in preferences should opt for the mixed logit model.

In light of the limitations of the conditional logit, the mixed logit is a very attractive model and has been increasingly implemented in applied research (Keane & Wasi 2013). The ML assumes that some or all of the estimated parameters in β are random and follow an assigned probability distribution, which most often is a standard normal. The mean coefficients of the attributes can be simply interpreted as the mean weights on utility of the different attributes that enter the model with respective standard deviations representing the estimated distribution of taste among the population. This assumption specifies that preferences are heterogeneous for that particular attribute. For example, if we assume that the price attribute is randomly normally distributed, then we are implying that different individuals will assign different weights to the effects of price on their utility, whereby the differences between individuals are captured in the standard deviations of the estimated mean coefficients. The researcher can specify the type of distribution that best suits their assumptions about the data, such as log-normal or normal. The *mixlogit* command in Stata accommodates both the log-normal and normal distributions which we also use to analyse our data in chapter three of this research paper.

The mixed logit choice probability is defined as:

$$P_{nit} = \int \left(\frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right) f(\beta) d\beta$$

(Train 2009, p. 138)

The above choice probabilities are calculated by taking draws from a mixing distribution $f(\beta)$, whereby "*[t]he mixed logit probability is a weighted average of the logit formula evaluated at different values of β, with the weights given by the density f (β). In the statistics literature, the weighted average of several functions is called a mixed function, and the density that provides the weights is called the mixing distribution. Mixed logit is a mixture of the logit function evaluated at different β's with f (β) as the mixing distribution*" (Train 2009, p. 135). This mixing distribution can be discrete, normal, triangular or alternatively uniform. Thus given a suitable choice of mixing distribution, the mixed logit is able to accommodate any form of utility maximising and non-maximising behaviour (Train 2009, p. 136). Assuming that the density of β is normal with mean *b* and covariance *W*, the choice probability becomes:

$$P_{nit} = \int \left(\frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right) \emptyset(\beta | b, W) d\beta$$

The researcher chooses an appropriate distribution for each of the respective attributes that enter the model and estimates *b* and *W*. Both *b* and *W* describe the density of β, and are the parameters of the distribution that can be denoted as θ (Train 2009, p. 136). For panel data, the mixed logit probability will be the unconditional probability of a sequence of observed choices over time periods *t*. For survey data this will be the sequence of choices made across different hypothetical choice scenarios. These preferences vary across decision makers but not between choices made by an individual decision maker. Thus the unconditional probability is derived by calculating the *product* of logit formulas to capture the sequences of choices *S* made by the decision-maker *n*. As we are calculating the product of the logit formula, we are using the product operator represented by Π:

$$S_n(\beta_n) = \prod_{t=1}^{T} \left[\frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \right]$$

The unconditional probability for the mixed logit using panel data is the integral of the above product of logit formulas. Calculations are made using simulation methods because the integral does not have a closed form solution unlike the standard logit formula. The choice probabilities will be evaluated over the values of β using a density function $f(\beta|\theta)$. As mentioned above, given a specific distributional assumption for the density of β, θ represents the underlying parameters of that distribution i.e. means and standard deviations. Thus we can calculate the choice probabilities by integrating the following function:

$$P_n(\theta) = \int S_n(\beta) f(\beta|\theta) d\beta$$

The estimation requires simulation methods and is founded on the maximum likelihood estimator which has a convergence criterion to the true value of the population (Greene 2008). We will be using Stata and the *mixlogit* command written by Hole (2007) to estimate the data. To fit the ML specification, the data is estimated by taking draws from an underlying distribution which the researcher can specify. This includes the type of distribution and the number of replications R in parameters β^r, where the rth draw is taken from the density distribution function $f(\beta|\theta)$. The *mixlogit* command accommodates both the normal and log-normal distributions. For example, the log-normal distribution may be specified if the researcher knows for certain that a particular parameter will not contain any negative values (Hole 2007). The analyst can compare whether a given distribution will improve the model fit by estimating the data under different assumptions. Following Train (2009, pp.144-145) and Hole (2007) the log-likelihood function for the model is $LL(\theta) = \sum_{n=1}^{N} lnP_n(\theta)$ and the simulated log-likelihood function below requires taking R draws from the assigned distribution of the parameters:

$$SLL(\theta) = \sum_{n=1}^{N} ln\left\{\frac{1}{R}\sum_{r=1}^{R} S_n(\beta^r)\right\}$$

Compared to the standard logit, which allows for a closed form expression of the integral containing the probability density distribution, the mixed logit estimation procedure requires decomposition of the density function of unobserved portion of utility $f(\varepsilon_n)$ into two parts. The first component contains the correlations in the data and the other follows an iid extreme value distribution like in the CL model. This first component can be assigned to any distribution and can therefore approximate any of the other choice model types. By decomposing the unobserved component of utility into two parts, the mixed logit is able to accommodate measurement of random taste variation to capture heterogeneity in the population.

We also note that the two main models (conditional logit and mixed logit) that have been outlined in this section accommodate post-estimation procedures. Namely, forecasting, calculations of elasticities of choice and willingness-to-pay estimates for service and product characteristics (Louviere et al. 2000, pp. 55-61). The next section review different approaches used to generate and evaluate candidate experimental designs including those which accommodate conditional logit and mixed logit models.

2.4 Choosing an Experimental Design

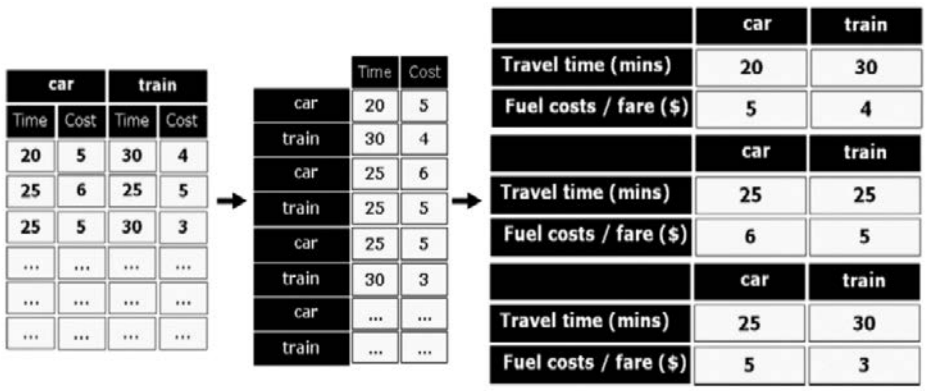
In this section we focus on two approaches in experimental design theory: orthogonal and *D*-efficient. In discussing these two alternatives, we note that there are associated trade-offs between these types of designs. The orthogonality property has often been considered the traditional and state-of-practice approach (Bliemer & Rose 2009). On the other hand, efficiency designs offer an attractive alternative orthogonal designs. Efficiency designs accommodate various discrete choice model forms at the survey design phase which can lead to important improvements in the precision of parameter estimates (Bliemer & Rose 2009). In addition, efficiency based designs reduce the sample size requirement needed to obtain robust parameter estimates. Below, we outline these two main approaches and their corresponding limitations. In doing so we place the most emphasis on *D*-efficient designs which we note on balance, offer the most flexible and attractive solution for the design of a discrete choice experiment. The techniques discussed in this section are applied using the Ngene software in the context of the designs of the pilot and final surveys.

Recall that the experimental design, defined by the design matrix $X = [X_n]$, contains the combinations of alternatives, attributes and their respective levels as assumed by the analyst (Ngene Manual, p. 89). The design matrix itself is used to populate the questionnaire that is presented to survey respondents (Bliemer & Rose 2009). In the previous sections we defined the experimental design of a stated choice experiment as representing the underlying combinations of characteristics that produce a given choice moment i.e. the survey question. The choice tasks that are presented to respondents over the course of a questionnaire, are derived from the underlying experimental design that the researcher manipulates at survey design stages. Choice experiments that do not require many alternatives, attributes and levels can be obtained using a full factorial design (Louviere et al. 2000, pp. 84-85).

The full factorial of an experimental design describes all possible combinations of attributes and all of their respective levels. However, full factorial designs are rarely used in practice because of the large set of choice situations that are required. In addition, the approach does not eliminate strictly dominant choice situations nor unrealistic ones. The choice situations produced by a full factorial design will depend on the number of attributes, attribute levels and alternatives. In practice, the final number of choice tasks is generally extracted from the full factorial design that contains the universal set of all possible profile combinations (Louviere et al. 2000, p. 90).We follow Bliemer and Rose (2009) and the Ngene software manual[34] to explain how to derive an efficient experimental design. As previously defined, during each survey question *t*, survey respondents *N* choose the preferred alternative out of *J* total number of alternatives over a total of *T* survey questions. The alternatives *J* are assigned K_j number of attributes and each attribute $k \in K_j$ has l_{jk} number of levels. The number of choice situations *T* created by a full factorial design will be: $T^{ff} = \prod_{j=1}^{J} \prod_{k=1}^{K} l_{jk}$.

Considering the above mathematical expression and product operators, the size of the design will be increasing in the number of attributes, levels and alternatives chosen by the researcher. For example, if we have a design with $J = 2$ alternatives, each alternative has 3 attributes and each attribute has 4 levels, the design will produce a total number of survey questions $T = (4*4*4)*(4*4*4) = 4{,}096$[35]. While a full factorial will be suitable for simple experiments, in practice, the total number of choice situations generated by a full factorial will produce too large a choice set for any one respondent to handle (Louviere et al. 2000, p. 90, Bliemer & Rose 2009). This concern can be overcome by implementing a fractional factorial design. There are different types of fractional factorial designs with corresponding assumptions and requirements that can be imposed by the analyst. The process of constructing the experimental design from a full factorial is presented under Figure 2.2. The diagram below shows that “car” and “bus” are the relevant labelled alternatives, whereas “time” and “cost” represent the attributes of these alternatives that vary respectively according to specified levels within the design i.e. magnitudes of the attributes.

Figure 2.2 – *Process of generating an experimental design*



[34] Ngene Manual, ChoiceMetrics, Version 1.1.1, p. 63, 2012

[35] *Ibid.*

Design Process: Full Factorial → Experimental Design → Choice Situations

(Bliemer & Rose 2009)

Firstly, to obtain good quality SP data, the survey design must reflect reality and equally must be able to isolate the individual contribution of each attribute on utility. Therefore each alternative within a given choice task must exhibit trade-offs in the eyes of the respondent and should vary sufficiently across choice tasks. For example, in every choice task, no one alterative should be strictly dominating the other (Bliemer & Rose 2011). Furthermore, the design must exhibit sufficient variation in attributes and levels throughout the survey questions (Louviere et al. 2000). To achieve the outlined requirements, researchers can rely on different approaches, each with their respective strengths and weaknesses. There are various fractional factorial designs, including orthogonal, random and different variants of efficiency based designs. In this paper we focus on *D*-efficient designs as they are the most widely accepted in efficient design theory compared with other measures (Bliemer & Rose 2009). Before discussing *D*-efficient designs in greater detail we consider a more traditional approach and associated limitations.

Prior to the emergence of efficient designs, orthogonal designs were commonly used in applied research because orthogonality is a well-known statistical property that is very desirable in linear models. A design is said to be orthogonal "*if it satisfies attribute level balance and all parameters are independently estimable.*"[36] Orthogonal designs are generated by imposing the property of orthogonality on the attributes contained in the columns of the design matrix (Bliemer & Rose 2009). The property was initially incorporated into SC designs because orthogonality has established statistical advantages found in linear regression models (Bliemer & Rose 2009) .

The variance-covariance ("VC") matrix of a linear regression model is given by $VC = \sigma^2[X'X]^{-1}$, where σ^2 is the model variance and X defines the matrix of attribute levels in the design or the matrix of data to be used in estimation (Bliemer & Rose 2009). When the matrix X is orthogonal, the elements of the VC matrix are minimised. This property ensures that the design does not exhibit multicollinearity and that the standard errors, (i.e. square roots of the sample variances) are minimised. When generating an orthogonal fractional factorial from the full factorial, orthogonality can

[36] Ngene Manual, p. 64, ChoiceMetrics, 2012

be maintained only to an extent because a fractional factorial design will create its own specific correlation structure within the design matrix.[37]

Orthogonality however, is not necessarily a desirable statistical property in the context of DCEs because discrete choice models are non-linear by definition (Bliemer & Rose 2009). Compared to *D*-efficient designs, the only clear advantage of using an orthogonal design is that it does not require *a priori* assumptions on the parameter values at design stages. Firstly, orthogonality does not ensure that a design excludes behaviourally implausible choice tasks. In these situations the design can be manually manipulated to remove implausible scenarios. However, these types of adjustments can distort the desired orthogonal correlation structure of the design (Bliemer & Rose 2009). In linear models, orthogonality is advantageous because the correlation structure prevents multi-collinearity and minimises the standard errors of parameter estimates. While this holds true for linear models, the property of orthogonality will most likely not be appropriate for the non-linear econometric models used at estimation stages, namely the mixed logit (Bliemer & Rose 2009). We explain below that the asymptotic variance-covariance matrix associated with the family of discrete choice models is calculated differently. Due to the limitations of the orthogonality property in discrete choice models, researchers have instead suggested the efficiency design approach may be an improvement (Bliemer & Rose 2009, Scarpa & Rose 2008, Quan et al. 2011). As a starting point, we provide an overview of the theory of efficient designs.

The "efficiency" of the experimental design refers to the expected standard errors of parameter estimates within the asymptotic variance-covariance matrix (Quan et al. 2011). By definition, data obtained using an efficient design will produce parameter estimates with the lowest possible standard errors. These designs are generated by incorporating the functional form of the specific econometric model at survey design stages. Hence, this method integrates the statistical properties of non-linear discrete choice models before any data is collected. There are several algorithms proposed in the literature that evaluate candidate designs based on restrictions imposed by the analyst. The algorithms that can be used to evaluate designs using the Ngene software are reviewed later in this section. The evaluation compares designs to locate the "best" design that will produce the smallest standard errors during the estimation

[37] Louviere et al. (2000) outline a detailed approach for deriving orthogonal fractional factorial designs.

stages (Quan et al. 2011). We will now outline how efficiency can be measured by the researcher when evaluating candidate designs. We also consider the types of assumptions that can be imposed on the experimental design.[38]

Efficiency in experimental design theory, refers to the standard errors contained within the asymptotic variance-covariance matrix of the model. The metric used to compare designs is the value of the determinant of the asymptotic variance-covariance matrix. Thus to calculate the relevant measure of the efficiency of a design, the analyst must firstly calculate the determinant of the asymptotic variance-covariance ("AVC") matrix denoted by Ω. Let us first define the AVC Ω_N matrix for a sample of N respondents. For the CL, the asymptotic-variance covariance matrix will be a K x K matrix that will generally depend on the experimental design $X = [X_n]$, vector of individual choices (or survey outcomes) $Y_N = \left[y_{jtn}\right]$ and parameter prior values assumed by the researcher. The design expected to generate the lowest comparable standard errors for parameter estimates should be the preferred design. Candidate designs are evaluated by comparing the *D*-statistic of each unique AVC matrix which assumes a single hypothetical respondent i.e. Ω_1. This approach produces a single statistic to facilitate the process of design comparisons and we explain this point further below.

As previously noted, the asymptotic variance-covariance matrix is not equivalent to that of a linear model's VC matrix. In this case, the maximum likelihood estimator ("MLE") of the AVC matrix for the CL model is the negative of the inverse of the expected Fisher information matrix (Scarpa & Rose 2008). The Fisher information matrix itself contains the second order derivatives of the log-likelihood function of the CL. The same procedure applies to the mixed logit model. The researcher must evaluate the information matrix that corresponds to the mixed logit model to achieve higher levels of efficiency for parameter estimates (Sándor & Wedel 2002). This procedure can be interpreted either as maximising information or as minimising the variance and standard errors of the parameter estimates in the AVC matrix. As shown by the function below, the AVC is derived either analytically or using simulation methods, by taking the negative inverse of the expected second derivatives

[38] The approach and adapted notation we present is largely taken from Chapter 7 of the Ngene Manual (2012) which summarizes the literature on efficient design methodology.

of the log-likelihood function of the discrete choice model's functional form that was outlined in the previous section (Bliemer & Rose 2009):

$$\text{AVC} = \Omega(\beta, x_{jtn}) = \left[E\left[I(\beta, x_{jtn})\right]\right]^{-1} = \left[-\frac{\partial^2 lnL}{\partial\beta\partial\beta'}\right]$$

With the model specific log-likelihood function L:

$$L(X, Y, \beta) = \sum_{n=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{J} y_{jtn} \log P_{jtn}\,(X, \beta)$$

The matrix is in effect populated by the researcher at survey design stages who selects the relevant attributes, levels and the prior values to be assigned to parameters in β. In other words, the AVC matrix Ω_N will depend on the underlying experimental design $X = [X_n]$, parameter values β and the outcomes (responses) of the survey i.e. choice indicator is $Y = \ y_{jtn}$ that can take on values of 0 or 1. In other words, y_{jtn} will be equal to 1 if respondent *n* chooses alternative *j* during survey question *t*, and equal to 0 otherwise. As a result, and contrary to orthogonal designs, researchers wishing to generate efficient designs must make assumptions about prior parameter values β to calculate the second order derivative of the log-likelihood function and to derive the AVC matrix. As the researcher does not know for certain the value of β, he or she must make an assumption on parameter value priors which we can denote as $\tilde{\beta}$. The assumptions the researcher makes on these prior parameter values will have a direct impact on the quality of the design and precision of final parameter estimates.

To achieve unbiased and accurate estimates during empirical analysis, the assumptions made during the design of the DCE must be as consistent as possible with true population parameter values i.e. the means and the variances of explanatory variables in the utility function (Bliemer & Rose 2011). Generally, when designing the choice experiment, researchers have at least some information on parameters, either from previous research or theoretical underpinnings, which will allow them to assign prior values to the model parameters (Scarpa & Rose 2008). Incorporating at least some information on the parameter values enables the researcher to better allocate the attribute levels within the design of the stated choice experiment (Bliemer & Rose 2009). This will deliver important efficiency gains in the design of the DCE, even if these are either small positive or small negative prior values (Scarpa & Rose 2008). There are two possible approaches to derive the AVC matrix; analytical or using *Monte Carlo* simulation techniques. The above log-likelihood function is the same for both the standard CL and ML models and as a result, only the resulting choice probabilities of the respective models P_{jtn} are different (ChoiceMetrics 2012, p. 90).

Researchers have also defined other criteria for choosing between designs that follow a similar approach to the *D*-error metric. These approaches impose different efficiency requirements, including but not limited to, minimising the standard error, accommodating willingness-to-pay estimation and minimising the sample size requirement. Regardless of which measure is chosen, any attempt to minimise the elements contained in the expected AVC matrix will minimise the expected asymptotic standard errors of the design (Bliemer & Rose 2009). For the purpose of our DCE, we focus on the most widely accepted measure found in current literature for evaluating efficiency designs.

This method requires the researcher to calculate a *D*-error estimate by taking the determinant of the expected AVC matrix that we outlined above (Scarpa & Rose 2008). Once the expected AVC matrix has been defined by the researcher the *D*-statistic can be calculated by assuming *N*=1 (i.e. a single respondent) because it is much easier to evaluate designs based on a single value (Scarpa & Rose 2008). The designs derived using the *D*-error criteria are referred to as *D*-efficient designs. The design with the smallest *D*-statistic out of all possible designs is referred to as a *D*-optimal design (Rose & Bliemer 2013). Due to the multitude of possible combinations of attribute levels for any given design it may not be possible to evaluate the *D*-error for every single candidate design of a given stated choice experiment. Hence it is the relative size of the *D*-error that will matter in the SC experiment. Following Bliemer and Rose (2009), we drop the *n* subscript as we assume a single hypothetical respondent and the *D*-error can then be defined as:

$$D\text{-error} = \det\left(\Omega(\, x_{jt}, \beta)\right)^{\frac{1}{K}}$$

As noted above, the AVC matrix is a *K* x *K* matrix and for the D-error to be independent of the size of the problem, the D-error is normalized by the power $\frac{1}{K}$ (Ngene Manual, p. 92). There are three variants of the *D*-error statistic most commonly used by researchers. These approaches differ in the *type* of knowledge that will be required to set the prior values. Equation (i) below is the D_z-error (*z* for “zero) which is calculated by assuming that the parameters in β are fixed and all have zero coefficient values i.e. that a particular attribute has zero effect on utility. Researchers have found that specifying non-zero parameters can significantly improve the efficiency of a design compared to using simple zero value priors (Huber & Zwerina 1996). Thus, alternatively researchers can obtain an estimate for a D_p-error (p for "prior") which requires that the parameter values in β be fixed, non-zero and known with certainty. The equation in (ii) below provides the variant for fixed non-zero priors.

On the other hand, researchers may want to address uncertainty over parameter priors. In this case, the alternative D_b-error (b for "Bayesian") should be used. Unlike other D-efficient designs, D_b-efficient designs require the researcher to assume a distribution that contains the parameters prior value means and variances i.e. the range of possible values that a given mean coefficient can take. For the Bayesian D-error computation, priors $\tilde{\beta}$ are assumed to be random variables with a joint probability density function $\emptyset$ that can follow either a normal or uniform distribution. Normal and uniform distributions have generally been the only type used in the literature so far (Ngene Manual, p. 92). This approach can effectively incorporate uncertainty over the assigned parameter prior values. These three alternative D-statistics are presented below with Ω_1 indicating an AVC matrix which assumes N = 1 respondents.

Variants of *D*-efficient designs:

(i) **Fixed zero prior values**: $D_z\text{-error} = \det\left(\Omega_1(\mathrm{X}, 0)\right)^{\frac{1}{K}}$

(ii) **Fixed non-zero prior values**: $D_p\text{-error} = \det\left(\Omega_1(\mathrm{X}, \beta)\right)^{\frac{1}{K}}$

(iii) **Bayesian prior values**: $D_b\text{-error} = \int \det\left(\Omega_1\left(\mathrm{X}, \tilde{\beta}\right)\right)^{\frac{1}{K}} \emptyset\left(\tilde{\beta}\middle|\theta\right) d\tilde{\beta}$

(Bliemer & Rose 2009)

The primary drawback of an efficient design is the need to assign prior values to parameters at the design stages. Any assumptions made during the design stage, including specification of priors, will influence the statistical efficiency of the design (Scarpa & Rose 2008, Quan et al. 2011). In this context, the generation of an efficient SC study must be adapted to the research objectives on a case by case basis (Quan et al. 2011). Bliemer and Rose (2009) test whether misspecification of priors leads to significant reductions in efficiency of a design. However, the authors find that even with significant differences between specified priors at design stages and true population parameter values, the D_p designs performed well and remained more efficient than orthogonal ones producing relatively more robust parameter estimates with comparably smaller standard errors. In most choice experiments, researchers will not know with certainty that the specified prior values are entirely accurate. After all, there would be no point in undertaking the SC study if parameter estimates were known with absolute certainty. At the same time, prior values play a central role in the generation of efficient SC designs (Bliemer et al. 2008).

Considering the concerns over prior value accuracy, researchers may wish to control for this uncertainty during the experimental design process. In this context, Bayesian designs offer a particularly attractive solution because they can accommodate the uncertainty about the true value of the parameter estimates in the form of a distributional assumption instead of imposing strict fixed prior values on the design. To evaluate the expected efficiency of candidate designs requires taking numerous draws from the underlying distributions of the priors which are usually assumed to follow normal or log-normal distributions (Bliemer et al. 2008). The process takes draws from the distributions of parameters using simulation methods. The statistic used to evaluate Bayesian D_b-efficient designs is provided in equation (iii). While all approaches require a search algorithm to evaluate different designs, to derive the D_b-error will require additional simulation procedures and computational burden.

There are several simulation procedures that can be used to evaluate D_b-efficiency which evaluated in detail by Bliemer et al. (2008). Irrespective of the simulation procedure used to find a design, the structured approach remains the same. Referring to the guidance of Bliemer et al. (2008), the approach for Bayesian approximation in experimental design generation is outlined as follows:

(i) Take *R* draws from the specified random distribution of parameter prior values to obtain a possible parameter value ;

(ii) the *D*-error is calculated for each of the parameter values β_k;

(iii) the D_b-error is given by the average *D*-error calculated over the different parameter values.

The simulation procedures that take draws from the underlying distributions of model parameters tend to differ in the way that draws are taken from the prior distributions; usually either systematic or random draws. Equally, the convergence rate of these simulation procedures will be different. Bliemer et al. (2008) test the performance of available simulation procedures by considering their ability to approximate the D_b-error to the true efficiency of the design and their speed. The different simulation methods are tested using the CL, but the results and recommendations presented by the authors are also valid for the mixed logit model. In the literature, researchers have mostly used *Pseudo-Random Montel Carlo* ("PMC") simulation to derive D_b-efficient designs. The PMC simulation procedure takes *R* independent draws for each of the β_k parameters from their assumed prior distribution. The researcher specifies the number of draws that should be replicated. Then the D_b-error is computed for each of the draws and then the average D_b-error can be calculated.

The process can be described as taking draws $\tilde{\beta}^{(r)} = \left[\tilde{\beta}_1^{(r)}, \dots, \tilde{\beta}_K^{(r)}\right], r = 1, \dots, R$, from the prior random distribution where the probability density function has *K* number of parameters $\emptyset_k\left(\tilde{\beta}_k \middle| \theta_k\right)$. PMC draws are taken in a random fashion. This is contrary to the deterministic draws of an intelligent structure attributed to *quasi random Monte Carlo* methods. Randomness found in PMC methods may not always be a desirable property making the alternative quasi random sequences an attractive option. In their paper, Bliemer et al. (2008) find that PMC methods perform much worse in finding designs close to their true efficiency value than the alternative *quasi random Monte Carlo* methods which include both *Halton* and *Sobol* sequences.

Regardless of the sequence type, reducing the number of draws will always reduce the accuracy of the model efficiency estimates. Compared to other procedures, Halton and Sobol sequences require fewer draws to achieve convergence to the true value of the design's *D*-efficiency. The primary difference is that Sobol sequences offer more coverage of the different *k* number of attributes and their dimensions. Thus although they are functionally similar, the Sobol sequence will tend to converge to the true value of the efficiency measure quicker than when using Halton draws. The distinguishing features of the different sequences are outlined in detail in the paper of Bliemer et al. (2008). The Halton and Sobol sequences, among others, are available as part of the Ngene software.

To address some further concerns over uncertainty, Rose et al. (2009) propose the *model averaging approach* for evaluating and selecting experimental designs.[39] This approach accommodates a variety of designs in the calculation of the average *D*-error, namely the preferred D_b-error incorporating Bayesian approximation. Using the Ngene software, this can be achieved by specifying several slightly different utility functions within the programming syntax. For example, the researcher may not know whether they will want to use a basic conditional logit or mixed logit model to estimate

[39] The model averaging approach for creating efficient designs is supported by Ngene.

their data. The model averaging approach accommodates this uncertainty by allowing the researcher to specify both these models when evaluating candidate designs. The efficiency measure will be then be the average *D*-error of the design. The researcher can also assign weights to each of the utility specifications on the basis of their most and least favoured utility specifications. The efficiency measure will then be the weighted average *D*-error of the design.

On the basis of the procedures described to this point, *D*-efficient designs can be computationally intensive to generate. However, they offer empirically attractive features. Let us now consider their impact on sample size requirements. Not only do efficient designs minimise standard errors of estimated parameters, there is also exists an inverse relationship between the sample size requirement and statistical efficiency of a design (Rose & Bliemer 2013). Traditional theories of sample size requirements for SC experiments are directly taken from general sampling theory not exclusive to DCEs (Louviere et al. 2000, pp. 261-265). We discuss sampling methods in the next section. Let us consider the Rose & Bliemer (2013) who suggest an alternative measure for efficiency based designs developed on the basis of samples size requirements.

The authors propose a new measure to determine the sample size requirement to achieve efficiency called the *S*-efficiency statistic. The authors note that efficiency based designs, *S*- or *D*-efficient, will produce not only more reliable coefficient estimates, but also compared to orthogonal fractional factorial designs will deliver on cost efficiency due to the smaller sample size requirements. This result is achieved regardless of the specific type of efficiency measure used because the approaches both rely on some form of minimisation of the determinant of the AVC matrix. Specifically, the *S*-efficiency measure is concerned with the relationship between standard errors of a given design and the representative sample size requirements to achieve statistically significant estimates. Just as for *D*-efficient designs, the first stage requires the calculation of the AVC matrix of a particular model.

Contrary to the *D*-efficiency criteria which focuses on the overall efficiency of a design, the *S*-efficiency incorporates the efficiency measure for individual parameter estimates. The process requires the calculation of the asymptotic *t*-ratios of each of the parameters considered for a particular design. With some algebraic manipulation, and assuming a vector of prior estimates $\overline{\beta}$ with a confidence interval of 95% Rose and Bliemer (2013) show that the sample size requirement is different for the different attribute parameters k. From the equation we can see that the sample size requirement will depend on the size of the predicted standard error with respect to the magnitude of the predicted coefficients priors. The lower bound sample size requirement is given by:

$$N_k \geq \left(\frac{1.96 * se_1(\bar{\beta}_k)}{(\bar{\beta}_k)}\right)^2.$$

The above equation illustrates how different types of parameters will require different optimal sample sizes to obtain statistically significant estimates, with more 'difficult' parameters requiring a relatively large sample size. Thus the *S*-error measure will focus on the most empirically challenging parameter in order to calculate the optimal sample size and can be expressed as follows: $S - \text{error} = \text{min.} \max_k \{N_k\}$.

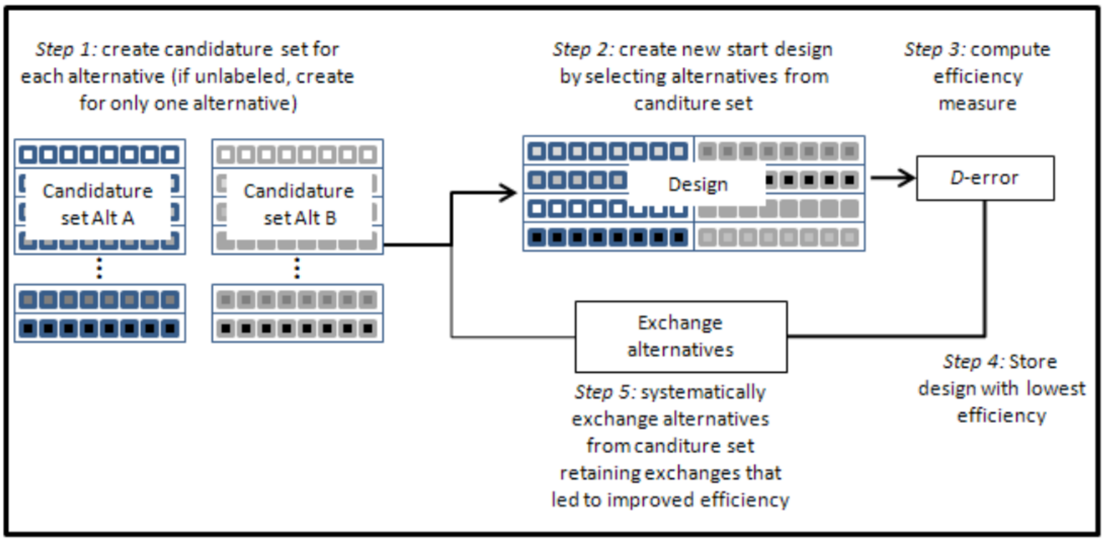
Sample size estimates based on this measure are directly related to the prior values attributed to the parameters. Hence this efficiency measure will equally be subject to the same uncertainties as for prior value assumptions. The Ngene software reports the *S*-statistic in conjunction to the other efficiency criteria that have been specified by the researcher. We note that as this efficiency measure is less conventional than others and we therefore favour a *D*-efficiency based design instead.

In this section we have presented the main differences between approaches in experimental design. In this context, we note that the evidence strongly suggests that it is worth investing in an efficient design. The efficiency focused approach enables the researcher to obtain more accurate parameter estimates, improves their forecasting ability and removes the requirement of a large sample size that would have otherwise been necessary to achieve statistically significant parameter estimates (Bliemer & Rose 2009). The next section focuses on other important considerations in survey design. We review choice of search algorithms (column vs row), aspects of labelled vs unlabelled choice experiments, sample size requirements, attribute level balance requirements and the optimal number of choice tasks to present to respondents.

2.5 Further Considerations in the Design of Discrete Choice Experiments

When searching for a design, using computer software like Ngene, the procedure relies on algorithms. In this particular context, once attributes have been specified by the researcher, algorithms are able to systematically search through all possible combinations of attributes and their levels by either swapping the rows *or* swapping columns of the design matrix (Bliemer & Rose 2009). The procedure allows the researcher to locate the designs that meet the criteria they have chosen to specify, including *D*-efficiency measures. Figure 2.3 below provides an example of a row swapping algorithm and the underlying process which evaluates designs based on a given efficiency measure.

Figure 2.3 – The Modified Federov Algorithm



(Quan et al. 2011)

In choosing what algorithm to use we refer to Quan et al. (2011) who assess the effectiveness of different algorithms that can be used while generating designs. The authors explore the main trade-offs between the two types of search algorithms used in SC studies, row versus column based swapping. Firstly, row based algorithms, as represented by the diagram above, offer speed advantages when used for the panel ML and CL models. Secondly, row based algorithms perform better when required to achieve *utility* level balance in a given design. On the other hand, column based algorithms are better suited for finding efficient designs that necessitate *attribute* level balance. However, column based algorithms do not easily satisfy the utility level balance property, and algorithms that swap rows are comparably less effective at achieving attribute level balance. In light of these differences between algorithms, the researcher must choose the algorithm that satisfies the properties that are more favourable for the particular DCE. In addition, researchers can use algorithms which are able to combine the two above objectives by systematically swapping rows and columns to derive different designs (Quan et al. 2011). We now consider issues related to labelled choice experiments.

When designing a choice experiment, the researcher has a number of different options in the way the survey is presented to respondents. One of these considerations is whether the experiment should be labelled or unlabelled (Louviere et al. 2000 p. 121, Bliemer & Rose 2009). Labelled choice experiments require alternatives to be described by brand or descriptive names, whereas unlabelled choice experiments *only* rely on the characteristics of the available alternatives. For example, a labelled choice experiment on preferences for ice-creams would present respondents with a choice between alternatives labelled as “Ben & Jerry’s” and “Hägen-Daz” or for automobile preferences alternatives could be labelled as “German car” and “American car”. In contrast, unlabelled designs describe alternatives by their attributes only. The alternatives can then be labelled as A, B, C…etc. or 1, 2, 3…etc.

Choosing between a labelled versus an unlabelled choice experiment is an important component of the design process. Brand names attributed to labelled alternatives signal additional information to respondents which can have indirect effects on choices they make between alternatives. Capturing these additional behavioural effects necessitates further empirical procedures when estimating the data. In practice, researchers have largely been ignoring the statistical effects associated with labelled choice experiments, partly due to the added estimation complexity (Louviere et al. 2000, p. 120, Doherty et al. 2013). However, the behavioural effects attributed to labelled choice experiments can affect the results at estimation stages and the design process must account for this. Labelled choice experiments will typically require additional parameter estimates and this increases the size of the SC study at design stages and also increases the degrees of freedom of the model (Quan et al. 2011).

In addition, the effects cannot be easily accounted for because they are unobservable to the researcher and are contained in the random component of the model (Louviere et al. 2000, p 120). For example, brand names can capture the emotional attachment that an individual has for a particular brand (Doherty et al. 2013). In this context, the respondent may not evaluate the all options presented to them in a given choice situation. When respondents do not fully evaluate options this is referred to as “attribute non-attendance” and this behaviour can influence the DCE. However this effect cannot be observed by the researcher. Labelled experiments may also cause respondents to infer information that has been omitted in the choice experiment and this may result in omitted variable bias (Louviere et al.2000, p. 121). We note that eye tracking technology in controlled experiments could prove useful in this context, but this approach is beyond the scope of this paper. There is experimental evidence on the impact of labelled choice experiments which we consider below.

Doherty et al. (2013) perform a DCE using the mixed logit model to explore whether labelled alternatives influence respondents' processing strategies when answering stated choice questions. In doing so, they are able to examine the effects of attribute non-attendance i.e. whether respondents fully evaluate the alternatives they are faced with. It is now widely recognized that individuals differ in how they process information as much as they differ in their preferences (Doherty et al. 2013). The authors explore the effects of labels by examining individuals' choice of recreational sight. The paper finds evidence that labels influence respondents' information processing strategies. The results show that a significant portion of survey participants had made their final choice by ignoring other alternatives. The experiment also showed that different demographic groups are shown to be influenced by labels to a different degree.

To avoid such issues, we propose an experiment design which omits the retailer name. Instead, we propose to provide respondents with realistic choice scenarios by including the attributes which reflect true retailer characteristics. This brings us to the issue of contextual realism. One of the main drawbacks of SP data are the associated difficulties in achieving contextual realism in the design of choice situations (Louviere et al. p 24). To achieve contextual realism, the researcher must choose *all* relevant attributes and associated levels that correspond to the market or product being studied. At estimation stages this will minimise the effects of omitted variable bias and improve the accuracy of results. Therefore market research and pilot studies are an important source of information for relevant attribute and level selection.

In preserving contextual realism, the researcher can also impose constraints on specific combinations of attributes and their levels presented to the respondents. This avoids nonsensical scenarios with combinations such as very low prices and the best possible quality. Even though such assumptions violate *attribute level balance*, this need not necessarily reduce design efficiency. The property of *attribute level balance* is the requirement that all attribute levels must appear an equal number of times for each attribute across the questionnaire (Bliemer & Rose 2009). Imposing this constraint on the design ensures that all the range of levels will be incorporated at estimation stages. However, attribute level balance may not produce the most efficient design (Bliemer & Rose 2009). This is a type of restriction reduces the number of candidate designs from which the researcher can choose their preferred option. In addition, this property will generally lead to designs with a relatively large number of choice tasks when creating relatively large[40] designs (Bliemer & Rose 2009). Removing the attribute level balance condition can actually improve the realism of a choice task. For example, for our SC experiment we want to ensure that the most expensive retailer would never have the lowest measure of quality. The flexibility of the Ngene software package enables the researcher to specify these types of constraints.

Let us consider the importance of *utility balance* in efficient designs following Huber & Zwerina (1996). Utility balance refers to the differences between utility values of the alternatives in a given survey question. This design property ensures that dominant alternatives that unquestionably generate the greatest amount of utility are not presented in a given choice task next to alternatives that would likely never be chosen. This is the type of nonsensical scenario we mentioned above. The most efficient design must exhibit some utility balance which accommodates just enough utility balance between alternatives to exhibit a degree of trade-off in the eyes of the respondents.

We previously discussed the implications of information processing in the context of labelled versus unlabelled choice experiments. The number of questions presented to respondents in a given survey can also influence individuals' engagement with the task at hand and resultant quality of responses (Hess et al. 2012, Bech et al. 2011). Respondent fatigue and boredom can induce individuals to not evaluate in full the characteristics attributed to different alternatives which may impact the parameter estimates. While it may be obvious that too many choice sets can reduce respondent attentiveness, the literature in this area is inconclusive and does not provide clear indication with regards to the "perfect" number of survey questions (Hess et al. 2012, Bech et al. 2011). In the literature, the number of choice tasks is usually in the range of 1 to 16 (Louviere et al. 2000, p. 261). To maximise response rates it is recommended to keep the number of survey questions to the lowest amount permitted by the design (Louviere et al. 2000, p. 261). Thus the key trade-offs to consider when choosing the number of questions to enter a design are (a) ensuring there are enough questions in order to capture a wide range of variation in preferences; (b) avoid deterring

[40] The design we use would be considered large because it has a total of 4 alternatives with 5 attributes that include 4 attributes with 4 levels and 1 with 5 levels versus a so-called small design which would only have 2 alternatives and around 3 attributes for example.

respondents by including too many questions and; (c) avoid increasing the likelihood of respondent fatigue by including too many questions; (d) respecting the degrees of freedom implied by the number of attributes entering the model.

Thus the aim of a smaller design has two objectives. Firstly to improve response rates, and secondly to reduce potential response bias that arises if respondents face too many questions. The process that determines the minimum amount of survey questions to achieve sufficient variation in the data, depends on the type of experimental design (Louviere et al. 2000, p. 259). We explained in the previous section that the full factorial increases with the number of alternatives, attributes and levels. Orthogonal designs typically require a larger number of choice sets compared to efficiency designs. This is partly because in order to achieve a design which maintains the orthogonality property will likely require many more choice sets than predicted by the degrees of freedom and attribute level balance (ChoiceMetrics 2012, p. 60). Using an orthogonal design will oftentimes require the researcher to split the full number of choice tasks amongst respondents. The optimal number of questions for an orthogonal design is derived by selecting the smallest main effects design from a full factorial. Irrespective of the design, the minimum number of choice tasks will be bounded by below by the number of degrees of freedom required and for the attribute level balance condition to hold (Bliemer & Rose 2009).

Recall that the property of attribute level balance is a condition that requires that attribute levels appear an equal number of times across the choice set. If the attributes differ in dimensions of levels this complicates the design procedure which must also achieve orthogonality in the correlation structure of the experimental design. In turn this tends to increase the number of questions in the candidate designs (Bliemer & Rose 2009). In practice it can be difficult to identify a design that incorporates these different features, including numbers of attributes and their different dimensions of levels in addition to the orthogonality condition. Efficient designs remove some of the restrictive properties of orthogonal designs and typically allow for a smaller number of survey questions. Efficiency designs also minimise the required sample size for a given experiment.

Rose and Bliemer (2013) highlight the lack of guidance on sample size requirements for stated choice experiments. The authors note that the literature suggests a sample size of 200 or 300 will typically achieve robust parameter estimates, however decisions on sample sizes are oftentimes made by researchers without a clear cut statistical argument. Following Rose and Bliemer (2013) there exists a relationship between sample size and standard errors of parameters in β that can be expressed as: $se_N(\bar{\beta}) = \frac{se_1(\bar{\beta})}{\sqrt{N}}$. Here $\bar{\beta}$ contains the means of the prior parameter estimates, the predicted vector of asymptotic standard errors of parameter estimates is given by $se_N(\bar{\beta})$ with a sample size N.

The above relationship represents a key feature of efficiency designs in the sense that "*an increase from 1 to 2 respondents will decrease the standard error by 29 percent, from 10 to 11 by 4.6 percent, from 100 to 101 by 0.49 percent, and from 1,000 to 1,001 respondents by 0.005 percent.*"[41] The equation above captures the reasons for investing in an efficient design: instead of increasing the number of respondents it is useful to invest in a more efficient designs because such an approach improves the accuracy of parameter estimates. We see that the marginal improvement of increased sample size on reductions in standard errors decreases steadily. An alternative approach to minimise the sample size requirement is to assign different blocks of questions to different groups of respondents (Rose & Bliemer 2013). However this comes at the expense of having sufficient variability in the collected data set.

For an efficient design, the degrees of freedom of the estimated parameters are also important. This is because the number of rows in the design matrix represents the number of choice situations that will enter the survey. In addition, the researcher may want to ensure that the aforementioned utility or attribute level balance properties hold within a design. Introducing these statistical properties into the design will further increase the dimensions of the design and thus the degrees of freedom. Following experimental design theory, we can express the minimum number of survey questions T as a function of maximum number of estimated parameters K and number of alternatives J that will be displayed in each survey question (Quan et al. 2011). In a choice experiment where respondents pick one option from a choice of J number of alternatives we will have $(J - 1)*T$ independent choice probabilities from each survey question (Quan et al. 2011). The lower bound maximum number of parameters will thus be $(J - 1)*T \geq K$, in addition to other constraints, such as attribute level balance.

Sample size requirements and optimal number of choice tasks are both related to each other. Each additional parameter to be estimated increases both the sample size

[41] Rose, J.M., Bliemer, M.C.J., "Sample size requirements for stated choice experiments", *Transportation*, Volume 40, Issue 5, 2013, pp 1021-1041

requirement and the necessary number of choice tasks. Furthermore, if a researcher is estimating interaction effects as well main effects from the explanatory variables, this will increase the number of degrees of freedom of the design. However, the importance of interaction effects and their prior values may not be known to the researcher at the design stages of a SC study. This presents an additional element of uncertainty associated with the generation of *D*-efficient designs. The next section reviews different methods for conducting surveys of human populations and associated biases which applies to the different methods.

2.6 Methods for Collecting Survey Data

This section looks at different ways researchers can implement surveys and draw a sample of the population of interest. In addition this section also evaluates different techniques to survey human populations including face-to-face interviews and web-based surveys. Typically the costs associated with performing a large survey are an essential consideration for many researchers, particularly PhD students. We therefore also assess likely costs, financial and time-related, associated with each of these methods. We conclude this section by outlining our proposed method for the distribution of our surveys as well as the collection of the data. In presenting our proposed approach, we also highlight the likely advantages as well as inherent biases attributed to this method. We briefly touch on ways to overcome such biases and how our approach can be improved upon in the context of a much greater budget.

Let us begin by looking at the different ways that a sample may be drawn from the population of interest, which in our case is the general population of grocery shoppers in the UK. We note that in drawing samples of the population there is the associated sampling error which arises simply because the process requires a sampling of the population (Stopher 2012, p. 270). However, this method allows the researcher to improve the robustness of parameter estimates compared to non-random sampling methods of data collection. When collecting data for the purposes of a study on a specific group of people, the collected sample must of course also be representative of that population of individuals or households (Stopher 2012, p. 68). There are a number of methods to achieve these two interrelated goals. Firstly, the sampling frame must be identified where the sampling frame itself represents the target population.

In the case of research which targets the *general* population, a suggested sampling frame can be a telephone directory (Stopher 2012, p. 266). However, this method has a number of downsides as the directory may contain out of date numbers.

More generally on the basis of the evidence, finding an adequate sampling frame of a human population is very difficult (Stopher 2012, p. 267). Instead, the researcher must define a sampling strategy when drawing a sample of a population, typically on the basis random sampling methods (Louviere et al. 2000, p. 262). Random sampling methods are advised in the context of modelling population parameters which are intended to represent the human population.[42]

The reasons for drawing random samples from a population in the context of surveys, is based on fundamental statistics theory. In other words, a sufficiently large and random sample of the population will be *unbiased* and tend to produce *robust* estimates for parameters corresponding to the population being studied. In practice, the most popular strategies are simple random sampling ("SRS" and exogenous stratified random sampling ("ESRS") (Rose & Bliemer 2013). The traditional approaches of SRS and ESRS look at minimising the sampling error that arises from random samples. SRS sampling typically requires a sampling frame and each individual in the sampling frame has an equal chance of being selected (Louviere et al. 2000, p. 262). ESRS sampling on the other hand, requires segmentation of individuals into mutually exclusive groups each representing a proportion of the population of interest (Louviere et al. 2000, p. 262). The researcher than draws from these segments until a sufficiently large sample size has been achieved.

As stated above, the main benefit of performing a random sample of the population is that it allows for the collection of an unbiased and representative data sample of the population and produces robust parameter estimates at estimation stages. This however requires a sufficiently large sample size. In surveys which rely on random sampling methods, required sample size is determined on the basis of which of these sampling methods is used. This is due to the relationship between parameter standard errors and sample size which captures the accuracy of parameter estimates (Louviere et al. 2010, p. 263). When undertaking random sampling methods, it is advisable to compute the minimum required sample size to avoid estimation problems following the collection of data.

We note however a number of difficulties associated with drawing a random sample from the general population. Firstly, there are issues in determining the relevant

[42] For a detailed evaluation of sampling methods please refer to Peter Stopher (2012) Chapter 13: Sample design and sampling, *Collecting, Managing, and Assessing Data Using Sample Surveys*, Cambridge University Press.

sampling frame when surveying the whole population. Secondly, due to the protection of individuals' personal information there are a number of challenges associated with the ability to reach the target audience of UK households and draw a sample from this population. Telephone directories have been used to carry out random samples of the population because individuals' information can be readily access. In addition random digit dialling using numbers randomly generated by a computer may be used to draw a random sample. These methods may however suffer from problems of non-response, may not necessarily produce a sample which is representative of sociodemographic characteristics in the population, can be time consuming and come at a significant cost to the researcher if they decide to purchase a sampling frame from which to draw from. For example, the sampling frame could be a panel of paid survey respondents, however this is also likely to lead to other types of biases.

The method of efficiency designs outlined above in Section 2.4, which is also our preferred approach, offers an alternative way to obtain unbiased parameter estimates of the population of interest. Efficiency based designs are fundamentally unrelated to sampling theory, however, they are able to deliver the same benefits in terms of producing robust parameter estimates. As such, efficiency based designs offer a way to overcome at least in part, some of biases which may arise when from *not* collecting a random sample of the population of interest. This may be the case when studying the general population and sampling frame does not actually exist or when using convenience sampling methods.

In this context, efficiency designs enable researchers to adopt more convenient data collection methods and still obtain robust parameter estimates. Although having a sense of required sample size in the context of efficiency based designs is important for estimated parameter robustness, using an efficient design, the researcher is able to obtain robust parameter estimates on the basis of smaller sample sizes compared to other approaches. PhD students in particular may opt for the most convenient approach when deciding how to sample the target population. This is referred to as "convenience samples" (Stopher 2012, p. 336). For example, the PhD student may approach and interview undergraduate students for the purposes of survey data collection. Stopher (2012) explains that this type of approach is perfectly acceptable to test certain types of hypotheses. Let us now consider the main methods for conducting surveys of the population.

Among the various methods used to collect survey data, face-to-face interviews were traditionally the most common method (Stopher 2012, p. 104).[43] In some surveys this type of technique can be essential if body language and respondent reactions are important in the context of the study. This method also typically requires a trained professional to carry out the interviews which represents an added cost to the researcher. In addition, Stopher (2012, pp. 105-106) explains that survey participants are likely to be put off by lengthy interviews, which can be exacerbated by human interviewers. Humans also can make mistakes in recording information during these interviews. Telephone surveys are also another way to collect survey data and this approach is relatively similar to the method of face-to-face interviews. The interviewer must read a script in the same way as for face-to-face interviews and record the responses. The same issues arise in the context of both methods in terms of human error and increased length of interview time (Stopher 2012, 109).

Another method which can be used to collect survey responses are postal surveys. Respondents must firstly be selected, say using one of the random sampling methods, then they are sent a survey via post and self-administer the survey (Stopher 2012, p. 107-108). Like all self-administered surveys, this type of survey requires the participant to be able to understand the survey questions they face. More generally, with the increase in the use of the internet, a very common approach among many practitioners are web-based surveys (Stopher, 2012, p. 104). This also represents a self-administered survey therefore respondents need to be able to understand the questions presented to them.

Some important benefits of computer based surveys are: they can ensure respondent anonymity making respondents more likely to be more truthful in their responses, they can prevent question non-response, minimise error and include automated prompts. For example, we explain below that we include a number of different prompts in our survey including asking respondents whether they are the main grocery shopper in the household. On the other hand, there are some limitations in using web-based surveys when conducting surveys of the human population. This issue arises because not all households have access to the internet and in some cases individuals

[43] In evaluating the methods for conducting surveys, we place emphasis on the suggestions and evidence presented in: Louviere, Hensher, Swait's (2010) book on *Stated Choice Methods* and Stopher's (2012) book *Collecting, Managing, and Assessing Data Using Sample Surveys.* Both of these books were published by Cambridge University Press.

may not possess the technological capabilities to be able to complete an online survey in the first place (Stopher 2012, p. 111). This may lead to more educated individuals participating in the survey than those from say, more disadvantaged backgrounds. Thus collecting survey data via web-based methods omits a proportion of the population from the sampling frame. Collecting data through a web-based survey will therefore likely lead to a biased sample containing younger, more educated and wealthier households. This drawback must be considered in the wider context of the costs and benefits associated with each of the approaches we have covered in this section.

In terms of “survey economics”, online surveys represent the most cost effective method in data collection techniques. However, there are a number of different trade-offs between different approaches, specifically, balancing accuracy and coverage with respect to the associated cost. The main cost categories in survey design are (Stopher 2012, (p. 356)):

- cost of drawing the sample;
- cost of building/ purchasing the sampling frame;
- cost of recruiting respondents;
- cost of surveying respondents; and
- cost of data processing, cleaning and checking.

Different alternatives in each of these categories represent their own cost to the researcher. Surveys can therefore be administered for a variety of costs. For example, surveys that do not require supervision, like online surveys, represent the cheapest alternative. Let us consider each other survey method in turn below (Stopher 2012, pp. 360-64). Postal surveys require a sampling frame to draw from, which is typically specifically designed for a survey. In this context, the sampling frame has to be purchased by the researcher which typically is very expensive. That is of course, if the sampling frame exists in the first place. Postal surveys require investments into the careful design of the survey because this method is one of self-administration which also comes at a cost.

The best quality method is the face-to-face interview and it comes with a hefty price tag. The evidence suggests that the cost of this methods is roughly $150-$500 per completed survey (Stopher 2012, p362). This represents a substantial cost to the researcher running the study. However, face-to-face interviews are associated with high response rates, the least sample bias and highest level of accuracy in information.

Different trade-offs are therefore inherent to all the available survey methods. On the basis of benefits and costs associated with the above approaches, in a perfect world with no budget constraints, the researcher would purchase/ acquire a sampling frame, draw a random sample from the sampling frame to reach the required sample size and then, of course, hire professional trained interviewers to perform face-to-face interviews. However, in the real-world, budgets are important and the *"art of survey design is one of trade-offs or compromise, and this is certainly the case in survey economics."*[44]

We acknowledge that online surveys may not be best suited for performing a study of the human population for the reasons outlined above. On balance however, online surveys have become commonly used for the collection of data and the most appealing aspect of online surveys are the cost advantages. We therefore propose to collect our survey data via an online survey platform and then test for sources of sample bias. We also favour "convenience" sampling as a means to collect the data, instead of drawing a random sample. This decision is also related to the time and cost associated with performing a random sample of the entire UK population.

Recall that of the main considerations we had when designing the discrete choice experiment were robust parameter estimates and the financial cost and time associated with collecting the survey data. In this context, we strongly argue that the approach we outline further below does not compromise on quality simply because we are collecting the survey data via web-based survey and adopt a convenience sampling approach. It is important to emphasise that we invested heavily in designing a survey which prioritises robust parameter estimates. We revisit the issues associated with online surveys, in particular likely biases, in the context of the next chapter. In Section 3.2 we discuss the quality of the data we have collected, outline likely sources of bias and how we propose to control for them in our empirical analysis. The next part of this section outlines how we propose to collect our data via web-based survey.

2.7 Pilot Survey: Design

This section outlines a pilot survey designed on the basis of features of the UK groceries market. The structural design of the survey itself (number of questions, combinations of values etc.) relies on techniques based on *D*-efficient experimental design theory. In the previous sections we explained that efficient designs require

[44] Stopher, P., *Collecting, Managing, and Assessing Data Using Sample Surveys*, Cambridge University press, p. 356, 2012.

assumptions to be made on the prior values of the coefficients of the parameters entering the design. The researcher must therefore establish *a priori* some understanding of the parameters being tested. A pilot study is therefore a useful tool to obtain prior information on attributes of interest. We emphasise that the aim is not to obtain precise parameter estimates. Instead, the goal is to roughly estimate the weight that individuals place on the different attributes entering the design. Section 2.4 presented evidence that in the context of *D*-efficient designs, inclusion of prior values when evaluating candidate designs, leads to important improvements in efficiency of the chosen design at the end of the process (ChoiceMetrics 2012, pp. 99-100).[45]

In addition, the pilot also enables the researcher to obtain feedback from participants on how to improve the aesthetics of the survey itself. The results of the pilot survey are presented in the next section. This section firstly explains every stage of the pilot survey design. We discuss the pilot survey's target sample size, the algorithm used to evaluate designs, how the attributes were chosen to enter the design, how the attribute levels were chosen, present the attributes and levels selected to enter the design and how the design analysis was performed in Ngene.

In terms of the target pilot survey sample size, at the outset we did not anticipate to collect a large number of responses. Typically, to obtain meaningful results, the minimum recommended number of responses is 100 up to a few hundred responses per survey, including in the context of pilot studies (Stopher 2012, p. 256). Our aim was to collect 25-30 responses given the resources available. Therefore, we acknowledge that the results presented in the next section are likely to be representative of a group of individuals *not* the population of UK households. With greater resources we would have sought to have a pilot survey with a sample size of at least 100 participants. Even with a small number of responses, we are able to achieve the main goal of the pilot survey which is outlined above. In other words, the information we collect enables us to get a sense of the weight that individuals place on different grocery retailer attributes that we can use in the form of prior values to generate the final design of the survey. We also rely on Bayesian methods which accommodate uncertainty about prior values

[45] Recall that in Section 2.4 we explained that specification of zero priors is the least effective way to draw from a set of candidate designs as this will not produce the most efficient design for that specific set of attributes and levels. We noted that on the basis of the evidence in the literature, specifying small negative or positive prior values represents a better way to determine the most efficient design.

assumed at the design stages. We explain this further when outlining the design of the final survey.

An important consideration in the design of a survey is that the researcher carrying out the survey and consumers in the population of interest are unlikely to share common preferences. In this context, they are likely to place different weight on different product attributes (Louviere et al. 2000, p. 257). To reconcile this problem, a qualitative assessment can help define the attributes that are likely to affect utility in the target population. We undertook qualitative research into the UK groceries market to inform the design of the stated choice experiment, namely looking at the findings of the CC's market investigation into the sector. The CC looks at the importance of the retailer features to consumers by estimating demand using revealed preference data from over 13000 UK households obtained from Kantar. These individuals recorded details of their grocery shopping trips over a period of at least a few years. The competitive assessment looks at various aspects of the market including barriers to entry, consumer demand and defines the drivers of competition in the market. The CC finds that the important attributes to consumers, are prices of products, quality of products, range and number of products and level of service provided known as "PQRS" as well as the proximity. In the short run, these variables can be adjusted relatively easily by the retailers and are thus considered to be important components of the dynamics of competition in the market (Competition Commission 2008, p. 49).

The results of the CC (2008) investigation are corroborated by a comparative analysis between British and Spanish shoppers. Colomé and Serra (2000) analyse the relative importance of different attributes of supermarkets to compare British and Spanish shopper preferences. Respondents in the study rank a list of 9 attributes from most to least important. On average the top most important attributes chosen by UK consumers was the quality of products, convenience, available range and the prices of the products in that respective order. Colomé and Serra's (2000) result includes 'convenience' as a relevant characteristic for grocery retailer choice. This indicates that proximity, or driving distance from the store, will impact on consumer utility and the decision-making process.

Store proximity in the UK groceries investigation was outlined as an important consideration for defining the relevant geographic market. The CC (2008, p.26) paper presents results from a consumer satisfaction report on UK's grocery retail customers. The report finds that consumer satisfaction was highest with a greater number of competing stores within their proximity. A more interesting result perhaps, was that the level of satisfaction increased significantly in the presence of a small store located within 5 minutes of the individual. This improvement in satisfaction was unrelated to the brand of retailer to whom the store belonged to. We considered it sensible to measure this variable in terms of drive-time given that most UK consumers drive a car to go grocery shopping (Competition Commission 2008, p. 69). The results of the CC's report on distribution of stores in the UK informed our selection of the proximity attribute levels.

The consumer analysis performed by the Competition Commission (2008, p. 45) shows that within 20 minutes driving distance, around 85% of the UK population will have a choice of at least 4 different grocery stores. This is consistent with the chosen number of alternatives that enter the design. In addition, the empirical procedures undertaken in the CC investigation, used a maximum distance threshold of 20 minute drive time. The 20 minute threshold indicates the maximum amount a consumer is willing to travel to the grocery store (Competition Commission 2008, Appendix 4.2). The findings on geographic store locations and consumers' willingness to travel, suggest that with 4 options of grocery retailer, the respective proximity of the stores should all be within 20 minute driving distance. In context with the findings of the CC groceries market investigation, and to ensure sufficient variation in the data, we chose to assign four levels to the distance attribute of 5, 8, 12 and 17 minute drive times.

Having defined levels for proximity, quality, range and service we had to choose the values to assign to the basket price attribute levels. In its investigation report, the CC found that throughout a typical week, shoppers tended to do one big weekly shop with some additional low expenditure trips to the store. Thus following from this behavioural observation, our price attribute is expressed as a value representing the weekly average basket price that could be expected at a given retailer. We wanted to ensure that respondents would face realistic basket prices in terms of how much they actually spend on an average weekly shopping trip. In this context, we calculated actual price differentials between the retailers based on the most frequently purchased items by UK households. We then 'normalized' these figures to maintain contextual realism by reflecting the actual expenditure of an average UK household on groceries.

The prices for the respective basket of goods were collected using the retailers' online websites. Out of the "Big Four" grocery retailers, Morrison's did not offer online shopping at the time the price data was collected and therefore Morrison's was omitted from this study. Instead, we included retailer Waitrose, who offer online shopping to account for high income consumers with lower price sensitivity. The CC findings highlighted the fact that UK shoppers perceived the Big Four retailers, plus Waitrose, as "good substitutes" (Competition Commission 2008, Appendix 4.2). We also noted that given the recent branded product price match marketing strategies across the sector, prices of branded goods tend not to vary substantially between retailers. Therefore we collected data for the cheapest own-brand products to capture true price differentials between retailers.

The approach used to calculate grocery retailer prices is consistent with the Competition Commission report. The report (Competition Commission 2008) used the food items listed on the Consumer Price Index ("CPI") to determine the price differentials between retailers. Items that fall under the CPI list are those that UK consumers purchase the most frequently and are determined from the results of the annual Living Costs and Food Survey. The products that were included in our calculations for average weekly basket prices, are taken from the 2012 CPI list published by the Office for National Statistics ("ONS").[46] From the full 2012 CPI list, we used 129 products; food, non-alcoholic drink and staple household goods such as bin bags. We excluded items for which we could not find comparable products on the retailer websites; for example, for own brand pro-biotic drinks. For further details the reader is directed to Table A.2.1 in the appendix which lists all of the items and prices we included in our price calculations and also the items for which we could not find comparable products. Price data were collected using the grocery retailer websites over the course of two weeks during the month of January, 2013.

Once the price data was collected, the figures were adjusted to ensure contextual realism and to calculate the "real-world" price differentials between retailers. To achieve contextual realism, it was not sensible to present the actual the sum of prices of CPI items because these values would not be indicative of the actual average weekly shopping trip faced by a typical customer. Therefore we normalized the basket prices to the levels of typical weekly expenditure by an average household in the UK, while maintaining the price differentials between retailers. Firstly, we determined the "base" price using ONS data on average household expenditure on groceries. The ONS report

[46] http://www.ons.gov.uk/ons/rel/cpi/cpi-rpi-basket/2012/cpi-and-rpi-basket-of-goods-and-services---2012.pdf

on households' expenditure states that the average UK household had 2.3 people with an average weekly expenditure of £53.40 on food items and non-alcoholic drinks.[45]

In this context, one of the four retailers was assigned the base price of £53.40. We chose Tesco's profile to indicate the base price because it was the most "mid-priced" retailer out of the 4 included. The results indicate that Waitrose is significantly more expensive than the other three retailers with a sum of prices equal to £305.57. Second most expensive is Sainsbury with £234.34, Tesco at £216.54 and Asda with £201.32. Using these sums of CPI listed product prices, we calculated the price differences between retailers in the form of percentages. In turn, these percentages were used to calculate 3 additional price levels. As we included both food and non-food items on our product list, we needed to adjust the average basket price to reflect the addition of these products. The non-food items, as a proportion of the sum of the total 129 product prices, account for 7%, 5%, 5% and 6% of the total value for Waitrose, Sainsbury, Tesco and Asda respectively. The basket prices were augmented to account for non-food items using these percentages. The procedure enabled us to derive the 4 different price levels based on "real-world" price differences between leading UK grocery retailers as presented under Table 2.1 above (decimal places were omitted). Following this approach, the total basket prices are both reflective of the typical weekly expenditure by UK households, and account for actual price differentials between the main grocery retailers active in the UK market.

Having calculated the price levels, we used these values to compute the attribute levels for the loyalty scheme discounts for two of the four retailers. At the time of our study, only Tesco and Sainsbury offered consumers the ability to collect points using a loyalty scheme. Our calculations included the double or treble your reward promotion that has been regularly used by the retailers to provide some variability in the levels of the discount. The annual discount was calculated using the value of the average weekly basket price over the period of a year. The calculation we performed to derive the discount values followed the same loyalty point formula used by Tesco and Sainsbury to reward their customers at the time the experiment was carried out. For both retailers, the consumer receives £1 discount for every £100 they spend in store or online and the loyalty points can add up to a sizeable sum over the year. Recall that the discount collected via the loyalty schemes can be spent on a variety of products and activities such as travel and festivals. Further details on the loyalty scheme structures the reader is directed to the first chapter of this thesis.

On the basis of the design considerations and the qualitative evidence outlined above, we chose a total of 6 attributes each with 4 levels to enter the experimental design. The loyalty discount represents the main variable of interest in the context of this experiment and is also an important part of several retailers' business strategies in the UK. These chosen attributes capture fully the most important features of the UK groceries sector and also represent important drivers of competition. We note that we chose to present 4 alternatives in each survey question because the vast majority of consumers have 4 grocery stores to choose from on any given shopping trip (see above). These are summarized under Table 2.1 below.

Table 2.1 - Attributes and levels in the experimental design

Attribute	Attribute Description	Attribute Levels
Basket Price	The price of an average weekly shopping basket of goods including food, non-alcoholic beverages and basic nondurable household items.	£53, £56, £61, £81
Loyalty Scheme Discount	The annual loyalty discount the average consumer can expect to receive.	£0, £29, £32, £58, £63, £117
Travel Time to Store	Store location based on driving time to the store in minutes.	5, 8, 12, 17
Product Quality	The overall level of own-brand only product quality the consumer can expect in their shopping basket.	Low, Medium, High, Very High
Product Range	The extent of product range in store in terms of product variety both within and across product categories.	Low, Medium, High, Very High
Service Quality	The quality of service a customer can expect in store in terms of staff politeness, queuing times, cleanliness etc.	Low, Medium, High, Very High

The final set of attributes entering the design are the *Average Basket Price, Loyalty Scheme Discount, Store Proximity, Service Quality, Product Quality* and *Product Range*. The attributes representing quality, range and service were assigned 4 levels to allow for sufficient variation within the experimental design, and to account for true market characteristics. The survey also presents respondents with a description of each of the above attributes and provides some examples to help with the interpretation of the attributes. As noted in the previous section, there are risks associated with self-administered surveys because individuals do not have the assistance of an interviewer who is able to explain the question to the survey respondent. It is therefore recommended to take extra care when designing a survey which requires self-administration.

The survey presented to respondents therefore applies descriptive language *not*: "Quality of Service: Very High". The survey contains the type of description to help the participant understand how to interpret low or very high level of service. For example, in the survey, service levels are defined as follows: "*Standard of Service: Overall friendliness and helpfulness of staff, check out waiting times, type of returns policy, cleanliness of the store, availability of parking spaces and overall shopping experience.*" The interpretation of these attributes and their levels was also facilitated by a description in the introduction of the survey. Survey participants were aware that the survey was implemented to study the UK groceries sector and that it was designed on the basis of features of this market. Therefore, participants can rely on their experiences of shopping across different retailers to gauge the significance (or insignificance) of these variables when choosing between retailers. We acknowledge however, that this does leave room for interpretation. In the empirical results chapter, we evaluate the estimates obtained for the qualitative attributes in terms of the weight we can place on their role in driving households' choice of retailer.

Let us also consider store size, which in the case of the CC model, enters as the actual size of the store in terms of square feet and is defined as an important driver of store choice. We considered that displaying square footage of a store to survey respondents may require some degree of interpretation. Instead, we use the 'range' attribute as an indicator of store size. This reflects real world markets where stores with low levels of range of products and product categories tend to be small in size, while stores with an extensive range of products and product categories tend to be larger supermarkets or so-called "hypermarkets". Similar detailed descriptions were provided to respondents and these can be found in the Appendix. Having defining the relevant parameters to enter the design, we now consider the experimental design evaluation process we undertook using the Ngene software.

In Sections 2.4 and 2.5 we explained the theory of experimental design and outlined the flexibility afforded to the researcher relying on this approach. The list of options available to the research when determining the experimental design to use for the study, which are also as accommodated by the Ngene software) can be summarized as follows:

i. type of design e.g. *D*-efficient, WTP, *S*-efficient;
ii. whether attributes are generic or alternative-specific;
iii. number of alternatives to present to respondents;
iv. attributes and levels to describe alternatives;
v. constraints on combinations of alternatives, attributes and levels;
vi. search algorithms to evaluate designs (e.g. column vs row based);
vii. the number of rows to include in the design matrix (i.e. number of survey questions);
viii. choice of using dummy or effects coding for qualitative variables;
ix. parameter prior values and distributional assumptions;
x. interaction effects;
xi. Bayesian approximation; and
xii. use of the model averaging approach.

We chose to specify generic alternatives within our design where respondents are presented with a choice of "Retailer A", "Retailer B"…etc. To avoid confusion during the design generation process, we coded the 4 alternatives as T, S, A, W within the Ngene syntax code. These labels enabled us to impose constraints on the price, quality and discount combinations. The constraints imposed on the design aim to maintain a degree of realism in terms of the grocery retailer profiles in the real world. Firstly, the most expensive retailer was assumed to never exhibit the lowest level of product quality. Equally the cheapest alternative was specified not to be combined with a very high level of product quality. In addition we assigned constraints to ensure that the different discounts correspond to their respective basket prices. In addition, the design was constrained to include two alternatives without a loyalty discount to account for the two retailers who do not offer loyalty schemes

When evaluating the efficiency of designs, the researcher can choose between various algorithms which systematically search for different designs by adjusting and alternating the combinations of levels of attributes within the design matrix (Scarpa &

Rose 2008). One can then compare the range of designs and their respective *D*-errors to find the lowest comparable error estimate. Considering the different properties of search algorithms, in the design of our survey we argue that the most suitable is the most widely used row swapping algorithm; the *Modified Federov algorithm* (Scarpa & Rose 2008). We opted for this algorithm due to the speed and statistical properties associated with this type of approach. The row swapping algorithm draws choice sets from a full factorial or fractional factorial and calculates the *D*-errors for each design it constructs. The process is repeated until a specified "stop" criteria has been achieved, for example, when a certain number of iterations has been performed (Scarpa & Rose 2008).

The software also allows the researcher to specific the econometric model which will be used to estimate the data. An accurate specification improves the efficiency of the design. We previously explained that our preferred model is the mixed logit. However we also noted the difficulties in assuming a mixed logit specification during the entire design evaluation process due to the computational burden. This problem is exacerbated when handling large designs with multiple alternatives and levels, such as our DCE. Following the advice of the Ngene software creators, we specified the CL for our pilot design given the common functional form between the CL and ML.[47] In any case, we did not plan to run the pilot data using a mixed logit specification because due to the small sample size this would not have been informative.

To evaluate the various designs, Ngene requires the researcher to specify utility functions for each alternative to define the attributes, levels and corresponding prior parameter values. Recall that the utility function can contain either generic of alternative-specific parameters. Therefore the utility function provides the basis for the design of the experiment. The software then derives the AVC matrix by assuming a single respondent which is used to calculate the *D*-error to evaluate the efficiency of candidate designs. The software systematically evaluates candidate designs subject to the search criteria outlined in the program command syntax and saves the ones that are found to be the most efficient, including for example, D_b-efficient designs that are found to have the smallest comparable D_b-error.

[47] The Ngene forum provided the relevant platform to obtain guidance from the creators of the software and can be accessed here: http://www.choice-metrics.com/forum.html

Instead of requiring the researcher to manually define a design matrix, the researcher specifies a representative utility function for each alternative that will be presented to respondents. The program automatically converts the utility functions into a design matrix that is evaluated using a search algorithm. This approach enables the researcher to assign specific attributes and levels to a given alternative, whereby the utility function contains the attributes and their levels that populate the experimental design. The Ngene syntaxes used to generate the pilot and main surveys can be found in Appendix.

To generate a D_p-efficient experimental design for the pilot study, we could either assume zero value coefficients, small negative or small positive coefficient values for the mean coefficient prior values $ß^k$. As discussed in Section 2.4, researchers have shown that by specifying non-zero priors at the design stage can yield gains in design efficiency compared to assuming zero prior values (Huber & Zwerina 1996, Bliemer & Rose 2009). Hence we opted to specify small prior values instead of assuming zeros throughout to increase the efficiency of the design compared to the zero value benchmark. Following the literature, we applied fundamental economic theory and logical reasoning to determine the signs of the coefficients (ChoiceMetrics 2012, pp.99-100).

For example, the price coefficient was allocated a small negative value by considering basic economic theory that suggests that increasing the price level will decrease utility of consumers holding all else equal. Based on our qualitative research the price of a retailer is a key driver of competition in the market. Hence, the price attribute was assumed to have the greatest weight on utility compared with the other variables. Similarly for the proximity attribute, we assigned a small negative mean coefficient value to account for the opportunity cost associated with increased travel time to the grocery store. On the other hand, the discount coefficient was assigned a small positive coefficient because a discount will increase consumer utility. Following the Ngene forum advice, a “small coefficient” can have a value equal or less than 0.01. In light of the lack of information on the intensity of the effects of the different attributes, we opted for even smaller coefficients of ‘0.0001’ for some of the variables. The parameter priors that were assumed to generate the pilot design are displayed under Table 2.2. The negative values for the qualitative variables represent the effects coding applied to the design of the survey because the value of the omitted coefficient is actually equal to the negative of the sum of the other coefficient values. We explain this further below.

Table 2.2 - ***Pilot study parameter priors***

Attribute k	Prior Value $\tilde{\beta}^k$
Basket Price	-0.01
Loyalty Discount	0.0001
Store Proximity	-0.001
Product Quality (Medium, High, Very High)	-0.0002, -0.0001, 0.0001
Product Range (Medium, High, Very High)	-0.0002, -0.0001, 0.0001
Service Quality (Medium, High, Very High)	-0.0002, -0.0001, 0.0001

Another option during the design process was either to use dummy coding or effects coding for the qualitative variables quality, range and service. Both dummy and effects coding require the researcher to omit one of the levels e.g. low quality. The main difference between the coding schemes is that, unlike dummy coding, effects coding has an additional level with the value '-1' assigned to the reference level which offers more variation in the data (Bech & Gyrd-Hansen 2005). When there are a lot of qualitative variables with several levels within the experimental design, the design matrix will contain a lot of '0' and '1' values that represent the base level for a given variable. This aspect of SC data can complicate the maximisation procedures at the estimation stage due to insufficient variation and equally can cause difficulty in finding an efficient design (ChoiceMetrics 2012, p. 124). Therefore, compared with dummy coded variables, effects coding can overcome the problems of non-convergence when using maximum likelihood estimation.

While dummy coding is more widely used, effects coding has its benefits in discrete choice modelling (Bech & Gyrd-Hansen 2005). Unlike for dummy coded variables, when using effects coding, the reference level, i.e. the omitted level, has a coefficient equal to the negative of the sum of the other coefficient values. Both approaches are functionally equivalent and should produce the same coefficient estimates (Bech & Gyrd-Hansen 2005). At the pilot stage of the study we were uncertain whether dummy coding would produce sufficient variation in the data to achieve convergence via maximum likelihood estimation due to the many "0" and "1" values as a result of the 4 qualitative variables. Therefore to avoid non-convergence, we opted to use an effects coding specification for our pilot design. As we discuss below, when estimating our data we obtained the exact same coefficient values when using either dummy or effects coding. The effects coded qualitative variable in Table 2.2 omit the ‘Very High’ level for each of the variables. As explained above, the coefficient value of the omitted level will be equal to the negative sum of the other specified coefficients.

As well as the main effects, Ngene also allows the option of including interaction effects. These can also be specified in the program syntax. Section 2.4, briefly mentioned interaction effects. Norton et al. (2012), Ai and Norton (2003) and Greene (2010) provide important insights to the testing and interpretation of interaction terms in nonlinear models, such as the conditional logit and mixed logit. Only at estimation stages can the researcher verify whether interaction terms account for variation in preferences by applying the z-statistic to coefficient estimates. This basic statistical test can help inform the researcher whether these terms improve the goodness of fit of the model and indicate whether a particular interaction term accounts for substantial and statistically significant variation in preferences.

During the design process, particularly at the pilot stage, it is impossible to know for certain whether interaction terms capture any meaningful variation in preferences. Following advice on best practice in the Ngene forum, we did not include interaction terms for demographic variables at the design stage. Instead, we included some interactions between the main explanatory variables. Following logical reasoning, the variables selected as interaction effects were assumed to account for differences in their effects on utility. Our pilot design syntax included 9 interaction effects with small prior coefficients which can be found in the appendix. We interacted the price coefficient with proximity, discount, quality and service. We had to limit the number of interactions because increasing their number also increases the degrees of freedom of the design which in turn increases the sample size requirement.

Following the specification of design requirements and relevant constraints, the representative utility function specifications for each alternative were also defined. These can be located in the appendix of this paper. For the pilot we specified a ten row design so that each respondent would face 10 choice situations in order to keep the survey to a minimum length. This number of choice situations was sufficiently large to accommodate the degrees of freedom of the model as this was a basic conditional logit specification. The utility functions in the syntax included the attributes and levels displayed in Table 2.1 in addition to the parameter prior values displayed under Table 2.2. The four respective utility specifications that describe alternatives A, B, C and D can be outlined as follows:

(i) $U(A) = \beta^1 x_{1A\{56\}} + \beta^2 x_{2A\{0,29,58,117\}} + \beta^3 x_{3A\{5,8,12,17\}} + \beta^4 x_{4A\{0,1,2,3\}} + \beta^5 x_{5A\{0,1,2,3\}} + \beta^6 x_{6A\{0,1,2,3\}}$

(ii) $U(B) = \beta^1 x_{1A\{61\}} + \beta^2 x_{2A\{0,32,63\}} + \beta^3 x_{3A\{5,8,12,17\}} + \beta^4 x_{4A\{0,1,2,3\}} + \beta^5 x_{5A\{0,1,2,3\}} + \beta^6 x_{6A\{0,1,2,3\}}$

(iii) $U(C) = \beta^1 x_{1A\{53\}} + \beta^3 x_{3A\{5,8,12,17\}} + \beta^4 x_{4A\{0,1,2\}} + \beta^5 x_{5A\{0,1,2,3\}} + \beta^6 x_{6A\{0,1,2,3\}}$

(iv) $U(D) = \beta^1 x_{1A\{81\}} + \beta^3 x_{3A\{5,8,12,17\}} + \beta^4 x_{4A\{1,2,3\}} + \beta^5 x_{5A\{0,1,2,3\}} + \beta^6 x_{6A\{0,1,2,3\}}$

The above utility functions *U(-)* represent the utility that a consumer will obtain by choosing a particular alternative j from a total of $J = 4$ alternatives. These alternatives are labelled as A, B, C and D. The coefficients β^k are generic and indicate the coefficient (or utility weight) assigned to attribute k. The attributes and respective levels are described by x_{kl} i.e. x_1 indicates the price attribute, x_2 indicates the loyalty discount etc. Here the values in the subscript parentheses denoted by *A{…}*, indicate the different levels that a single attribute can take for that particular alternative. The letter "*A*" simply indicates that attributes are generic i.e. they appear in the first utility specification *U(A)* and also for the rest of the alternatives. These functions can be manipulated in the Ngene software to impose specific design constraints and assumptions. Namely to define which attributes and levels should be combined together to form a single alternative. For example, we were able to ensure that certain prices are combined with specific levels of quality and loyalty discounts. Therefore the third and fourth utility specifications above omit coefficient β^2 to indicate a loyalty discount equal to zero. This ensures that each survey question displays 2 retailer options who do not offer a loyalty scheme to reflect the real-world retail offer available on the market.

The design for our pilot questionnaire generated a D_p-error of 0.007128. As all designs are unique, the *D*-error is not an absolute measure and should therefore be compared against other designs generated for the same SC study. While the relative differences in *D*-errors between candidate designs is a fundamental consideration for efficiency designs, we have previously discussed other design features that are also important. For example, some of the designs we inspected during this process had relatively smaller *D*-errors than others, but the more "efficient" designs were not always exhibiting sufficient variation in attribute levels. Following a thorough inspection of the designs and their features, we opted for the design that had a comparably low predicted D_p-error and which met other design criteria, namely attribute level balance.

We chose to conduct our survey via the Qualtrics platform. The software has a number of advanced features. For example, we are able to embed prompts which a prospective survey participant has to review including the following:

- Are you over 18 years old?
- Are you responsible for carrying out most of the shopping on behalf of your household?

In addition, the software was instructed to automatically block individuals attempting to access the survey from outside of the UK. If the relevant checks are not met an automated response by the software blocks the survey and instructs the respondents that they are not eligible to be survey participants. We also did not let individuals proceed with the survey unless they had ticked a response in each of the questions. We randomized the order of the questions presented to respondents to avoid question order bias. For the purposes of the pilot survey, the survey link was emailed to a small number of UK grocery shoppers of variable ages, gender and occupations.[48] We were later able to discuss the respondents' experiences of taking the survey to improve on the final survey appearance. This approach represents a "convenience sample" which we evaluated in the context of Section 2.6 above. The Appendix contains print screens for sections of both the main and pilot questionnaires.

2.8 Pilot Survey: Results

This section considers the results of the pilot survey. In fitting the pilot data, it is common practice to rely on the conditional logit. We explained above that in the context of the *final* survey design, large designs are typically evaluated by assuming a conditional logit specification. We carried out the pilot survey over the course of a month between September and October 2013 and obtained responses from 26 individuals. Each respondent evaluated 4 alternatives throughout 10 survey questions

[48] For both the pilot and main surveys we wanted to avoid the significant costs associated with obtaining a completely random sample of UK grocery shoppers, thus we relied on several networks of individuals from university students and lecturers, social media platforms and company mailing lists.

which generated 1040 observations in total. After downloading the pilot survey results from Qualtrics, the data was cleaned and inputted into Stata.

The data we obtained from the pilot study provides an observation for each choice scenario that was evaluated by an individual respondent. Both the chosen alternatives and rejected alternatives are recorded as individual choice moments. These observations provide information on the relative effects of utility achieved from choosing a particular alternative. For the data to fit McFadden's conditional logit model, the Stata manual[49] recommends either the *mlogit, clogit* or *asclogit* as suitable commands. This explains why practitioners use the terms conditional logit and multinomial logit interchangeably. In principle, the Stata commands are identical because they produce the same outputs. The main difference between the commands is the command coding and the way that the data must be arranged in Stata's data browser. For example, the *asclogit* is an alternative-specific conditional logit model that requires the least amount of data manipulation when dealing with labelled alternatives. We ran both *clogit* and *asclogit* commands to ensure consistency and robustness of our results, and obtained identical results, while the results reported in this section have been obtained exclusively using the *clogit* command.

Within the Stata software, the qualitative variables for quality, range and service were recoded as dummy variables to enable us to measure the effects of these attributes at their different respective levels. We note that this approach increases the number of degrees of freedom but is far more informative from an empirical perspective. Instead of estimating a single "average" value for a qualitative variable, each level of the variable (low to very high) has its own corresponding mean coefficient value. Recall that we specified effects coding for the pilot design in case of estimation problems that arise from the many 0 and 1 values in the covariance matrix. Hence to test whether the results were the same for different coding structures, we compared the results for both dummy and effects coding. Using our data, we obtained the same estimates under both coding schemes. In light of the more prevalent use of dummy coding this is also our preferred approach.

As we explain in the next chapter, the results of a logistic regression produce coefficients that are in values of *log*(odds). These values can be interpreted as the relative weights assigned to the corresponding attributes in the underlying utility

[49] StataCorp.. *Stata 13 Base Reference Manual*. College Station, TX: Stata Press, 2013

function as set up by the analyst. In the literature, researchers oftentimes refer to the coefficient estimates as probabilities, but they are actually logit probabilities and should be interpreted accordingly. A basic logistic regression model can be expressed as $\boldsymbol{logit}(\boldsymbol{p})=\boldsymbol{log}(\boldsymbol{p}/(1-\boldsymbol{p})) = \boldsymbol{\beta}_0\mathbf{x}_0 +\ldots+\boldsymbol{\beta}_k\mathbf{x}_k$, where $\boldsymbol{p}$ is the overall probability of choosing a given retailer and $\boldsymbol{p}/(1-\boldsymbol{p})$ is in the form of an odds ratio.[50] For ease of interpretation, we have thus calculated the exponent of the coefficients using the post-estimation *or* command in *Stata 12*. The ***log***(odds) can be converted to odds by simply taking the exponent of the coefficients produced by Stata for attribute *k* i.e. calculating e^{β^k}. We include the exponent that has been calculated for the coefficient values under the columns titled O.R. i.e. odds ratios. These results are shown under Table 2.3 in a separate column. The below results cluster standard errors at the individual-level and therefore contain robust standard error estimates. Robust standard errors are presented below the mean coefficients in parentheses. We note however, given the small sample size we do not place emphasis on these results as they are highly unlikely to be representative of the UK population. As we explain in the previous section, these are merely indicative results.

[50] We follow the guidance of UCLA Statistical Consulting Group for interpreting coefficient values and corresponding odds ratios in logistic regression models which has been obtained from http://www.ats.ucla.edu/stat/mult_pkg/faq/general/odds_ratio.htm

Table 2.3 - Pilot survey results

	Model 1			Model 2	
Variable	**Coef.**	**|*z*-stat|**	**O.R.**	**Coef.**	**|*z*-stat|**
Price	-0.036*** (0.013)	2.75	0.964	0.012 (0.110)	0.11
Discount	0.01*** (0.003)	3.23	1.010	-0.105 (0.077)	1.36
Time	-0.081*** (0.025)	3.24	0.922	-0.234 (0.312)	0.75
Medium Quality	2.39*** (0.450)	5.3	10.881	18.313 (13.716)	1.34
High Quality	2.73*** (0.457)	5.98	15.330	20.432 (14.247)	1.34
Very High Quality	2.938*** (0.641)	4.59	18.885	22.92 (14.922)	1.54
Medium Range	0.744*** (0.243)	3.06	2.105	0.56* (0.305)	1.84
High Range	0.814*** (0.246)	3.31	2.256	0.793* (0.476)	1.66
Very High Range	1.101*** (0.340)	3.24	3.007	1.521*** (0.318)	4.79
Medium Service	0.481*** (0.244)	1.97	1.619	0.322 (2.189)	0.15
High Service	0.576 (0.347)	1.66	1.779	-0.820 (1.985)	0.41
Very High Service	0.772* (0.411)	1.88	2.165	3.203 (3.284)	0.98
Price*Discount				0.002 (0.001)	1.45
Price*Time				0.003 (0.005)	0.5
Time*Discount				-0.001 (0.001)	0.91
Price*Low Quality				0.293 (0.242)	1.21
Price*High Quality				0.293 (0.059)	0.58
Price*Very High Quality				-0.067 (0.062)	1.07
Price*Low Service				-0.0005 (0.035)	0.01
Price*High Service				0.014 (0.042)	0.33
Price*Very High Service				-0.051 (0.053)	0.95
Log-likelihood	-270.565			-262.059	
Nr. Respondents	26			26	
Nr. Observations	1040			1040	

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

Using the pilot data we were able to obtain statistically significant estimates at the 1% level for most of the first model's explanatory variables. The respective absolute *z*-values are indicative of the degree to which the respective attributes account for variation in grocery retailer choice. The pilot study estimates show a strong preference for higher levels of product quality compared to other retailer attributes. Product quality is more important to consumers than the quality of service, while the level of service is shown to be more important than the extent of product range and variety. Grocery retailer quality accounts for a substantial amount of variation in the data with a *z*-statistic of 5.46 attributed to a very high quality level. The results show that by offering high quality products in their store, a retailer will have much higher odds of being selected by consumers. More precisely, holding all else constant, if the product quality improves from low to very high, the odds that the consumer will choose a given retailer increase by well over 100%

This value is likely to be overstated because the calculated odds of 18.89 are unusually high for this type of measurement scale. As mentioned previously, the lack of observations is likely to have overstated the effect of product quality on store choice. On the other hand, the strong preference for quality could also be a reflection of the fact that the small sample of respondents could be from higher income households. The results show that the variables of price, time and discount are also important to consumers. The estimates show that consumers prefer lower shopping costs and value a shorter travel distance to the grocery store. In fact, the sample results show that consumers value proximity to the store location more highly than the actual basket cost of their shopping trip. This is shown by the negative coefficients for price and store proximity and a positive coefficient for the loyalty discount.

Looking at the results in Table 2.3, the discount variable has an odds ratio of 1.01. This result indicates that retailers that offer a loyalty scheme improve their odds of being selected compared to those who do not offer a loyalty scheme. More exactly, keeping all else constant, a £1 increase in the annual loyalty discount improves the odds of retailer selection by 1%.While the pilot stage of the choice experiment is essential, the pilot results must be approached with a degree of caution because they are restricted by relatively few observations compared to the complete study. Based on these results, it would be recommended that for D_p-efficient designs researchers should invest in a large enough pilot study to obtain more accurate prior values to enter the final design.

Recall that during the design stages we included interaction effects between some of the explanatory variables. We ran the model with and without interaction terms and evaluated the respective coefficient *z*-statistics and the results are presented in Table 2.3 above. The coefficient values were found not to be statistically significant. It is likely that our sample size was insufficient to capture these additional effects and the model produced some nonsensical results like a positive coefficient for price. In Section 2.4 we discussed how additional degrees of freedom can require additional data, through a greater sample size or additional survey questions. Having outlined the pilot survey results and the main design procedures which apply to both the pilot and main survey designs, we now proceed to the methodology underlying the final survey design.

2.9 Main Survey Design

The pilot coefficient estimates estimated using *Stata* and presented in the table above enter the final survey design as parameter priors. These priors are outlined in Table 2.4 below along with the assumed standard deviations. The population standard deviations of the respective attributes included in the design were approximated using the values of the standard errors of the coefficient estimates. We explain this approximation in further detail in this section. This section outlines a model averaging approach using Bayesian approximation to evaluate available designs subject to outlined uncertainties. We previously covered the main advantages of this approach, including the ability to account for model uncertainty and the precision of the parameter values obtained from a pilot study. We also apply the advice received from the creators of the Ngene software in the online forums and the Ngene manual (ChoiceMetrics 2012). They explain that larger designs like ours which also assume the mixed logit model should be derived using the conditional logit specification and then evaluated against the mixed logit in the final stage. This section concludes by presenting the sociodemographic questions which enter the final design of the survey.

The main survey design generation process took place over the course of many weeks to allow Ngene to evaluate as many potential designs as possible and to locate the smallest comparable *D*-error for the final questionnaire. The utility specifications for the final design are identical to those outlined in the previous section. Equally, as specified for the pilot design, the main survey syntax also used the Modified Federov Algorithm to evaluate different combinations in attributes and levels found in the design matrix. The key differences to the design generation process was the application of the model averaging approach and Bayesian approximation methods.

As explained throughout the discussion on experimental design generation in Section 2.4, Bayesian efficient designs require additional simulation procedures. Using the Bayesian approach the parameter priors are assumed to be random instead of fixed. The design evaluation process necessitates a predetermined number of draws to be taken in a random or systematic sequence from the underlying distribution. These draws are taken from each of the parameter distributions as specified by the researcher, where each distribution is defined by the mean coefficient value and its corresponding standard deviation. Following the Ngene Manual (p. 100), "*[a] Bayesian efficient design optimizes the expected efficiency of the design over a range of prior parameter values, thereby making it more robust to misspecifying the priors. Priors with a higher uncertainty should see this uncertainty reflected into a larger standard deviation or spread of its probability distribution.*"

In our model, each parameter prior $\widetilde{\beta_k}$ is assumed to follow a normal distribution with mean μ_k and standard deviation σ_k. The population standard deviation values in Table 2.5 have been approximated using the standard errors of the mean coefficient estimates obtained via the pilot study using Stata presented in Table 2.3. When considering point estimation of parameters, "*[t]he standard error of the estimate is the standard deviation of the sampling distribution of the statistic*" (Greene 2008, p. 1027). Recall that the standard error is concerned with the precision of parameter estimates (i.e. accuracy with respect to the true population mean) and decreases with increasing sample size. Whereas the standard deviation measures the dispersion of data in the population around the population mean and has no relationship with sample size.

In order to obtain a rough approximation of the population standard deviation σ we use the well-known relationship between sample size, the standard error of the parameter estimate and the population standard deviation: $\mathrm{S.E.}_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.[51] We follow this approach instead of randomly assigning values to the standard deviations. At this stage, the exact value of the standard deviation (and equally the value of the prior mean coefficient) will not be known with certainty and as noted above, the researcher can assign larger standard deviation prior values if there is uncertainty over the mean coefficient value. It follows that for the design generation process the standard deviation

[51] This approach has not been used to obtain a precise and unbiased estimate of the population standard deviation. We use this approach to avoid assigning random values to the distribution and acknowledge that our sample size of 26 is probably not sufficiently large for the formula to accurately predict the population standard deviation.

need not be a precise estimate of the true population standard deviation but rather should be used as a tool to indicate the degree of uncertainty of mean coefficient parameter priors. The table below specifies the values assigned to each of the distributions of parameter priors for every corresponding attribute.

Table 2.4 – Main survey parameter priors

Attribute		Assumed Prior: $\widetilde{\beta_k} \sim N\,(\mu_k, \sigma_k)$
Basket price		*N* (-0.04, 0.05)
Loyalty discount		*N* (0.01, 0.02)
Proximity		*N* (-0.08, 0.1)
Product Quality	Medium	*N* (2.39, 2.8)
	High	*N* (2.73, 2.7)
	Very High	*N* (2.94, 2.8)
Product Range	Medium	*N* (0.74, 1.38)
	High	*N* (0.81, 1.57)
	Very High	*N* (1.1, 1.71)
Service Quality	Medium	*N* (0.48, 1.34)
	High	*N* (0.58, 1.81)
	Very High	*N* (0.77, 1.48)

The D_b-error is calculated by Ngene firstly by drawing, R values from the random distribution of the prior parameter values as defined in the table above. Then, for each of these parameter values, the D-error is evaluated and an average D-error is computed over these values. Recall that different parameter values will produce an AVC matrix with its own D-error. However over a large number of draws, Bayesian approximation achieves convergence to the true D-efficiency of the experimental design. For the main survey syntax, the draws were specified to follow the *Halton Sequence*. Compared to its counterparts, the Halton sequence performs well because it generally requires fewer draws to converge to the true efficiency of a given design, as compared to PMC methods (Bliemer et al. 2008). Before deciding to use the Halton sequence, we also performed trials by specifying the Sobol sequence. Throughout the trial runs, both of the procedures produced relatively similar results. In the final syntax we specified a relatively large number of draws (40000) to be taken from the distribution. This was done because over a large number of draws the approximation converges to the true *D*-efficiency value of the experimental design irrespective of the sequence used (Bliemer et al. 2008).

In addition to prior value specification, we accounted for other uncertainties by using the model averaging approach. Following this approach, the researcher can specify several design specifications, i.e. models, with their respective assumptions. Then the researcher specifies weights for the models to indicate the degree of model preference/importance. The average *D*-error of the models is calculated using these weights. The main survey design syntax found in the appendix contains the three model types (M_1, M_2, and M_3) that were assumed for the model averaging approach. All three models were assumed to be CL with Bayesian priors as outlined in Table 2.5. We assigned different properties to these models to account for lack of information at design stages.

The differences between the models were the inclusion/exclusion of interaction terms and effects versus dummy coding. For example, the first model, M_1, assumed dummy coding and main effects only. The second model, M_2, included interaction terms (same as the ones used in the pilot design) and also used dummy coding. The interaction term priors were assigned small negative or positive values instead of using the pilot survey results.[52] The third and final model M_3 properties were assumed to be main-effects only with effects coded variables instead of dummy coded. Recall that effects coding may be important at data analysis stages if dummy coding hinders convergence via maximum likelihood estimation.

As before, Ngene calculates the individual AVC matrices for each of the models specified and calculates the *D*-error of each individual model. When using the model averaging approach, Ngene also displays the *D*-error adjusted by the weights assigned by the researcher. The weights assigned to the model types were introduced to ensure that the preferred model would be given the most weight in calculating the *D*-error. The corresponding weights assigned were 2:1.5:1 for models M_1:M_2:M_3 respectively. Model M_1 was given the greatest weight because it is our preferred main-effects dummy

[52] As a result of the small sample size, the estimation in Stata produced nonsensical values for the interaction term coefficients.

coded model that would also be used during the evaluation stage for the mixed logit. In addition to these considerations, the Ngene syntax required that we also specify the econometric model types for these three versions.

Using the model averaging method, we specified the CL model for all three model types. After allowing Ngene to run the syntax, we manually saved designs with the lowest D_b-errors. These stored designs were later evaluated by assuming the panel mixed logit model. The evaluation syntax that was used to instruct Ngene to evaluate the pre-saved designs can be found in the Appendix of this paper. We firstly attempted to derive a design that would produce ten survey questions i.e. ten rows in the survey design matrix. However, given the additional degrees of freedom required to estimate the mixed logit model, we needed to include an additional row within the design matrix. Therefore our final survey presented respondents with a total of 11 survey questions. When we evaluated the designs against the ML (random parameter) specification we also had to select the number of hypothetical/simulated respondents (ChoiceMetrics 2012, p. 112). With random parameters, preferences are represented by individual-level coefficients β_n^k that are drawn from a particular distribution.

Hence to calculate the *D*-efficiency of a design that assumes a mixed logit model with random parameters, the process requires the analyst to take a specified number of draws from the distribution for each hypothetical respondent and then to calculate an average *D*-error for the design. Larger numbers of draws and greater numbers of hypothetical respondents produce more accurate approximations of the efficiency measure at an additional cost of increased computation time. We used a relatively large number of N = 1000 respondents with 300, 500 and 1000 Halton draws to evaluate the efficiency of the candidate designs. The specific *D*-error estimates of our chosen design can be found under the table below.

The bottom row in Table 2.6 below indicates the final *D*-error of our chosen design when it was tested against the mixed logit model specification. The figure in the brackets indicates the number of Halton draws. Recall that the *D*-error is determined by the individual design and assumptions imposed, hence the *D*-error values can only be compared between designs of a given choice experiment.

Table 2.5 – *D*-errors of the main survey design

Testing the *D*-error of the chosen experimental design				
Model	**D_b-error** (H=40000)	**D_p-error** (H=300)	**D_p-error** (H=500)	**D_p-error** (H=1000)
M_1	0.242081			
M_2	0.053306			
M_3	0.355818			
Unweighted Total Error **(M_1, M_2, M_3)**	0.651205			
Average Weighted Error **(M_1, M_2, M_3)**	0.919939			
Panel Mixed Logit **(*N*= 1000)**		0.49089	0.49808	0.49686

Next we present the sociodemographic questions we propose to ask the survey respondents and look at the ways these questions should be formulated in the context of a survey. In preparing these questions we noted that there are a number of known issues related to the truthful elicitation of responses, in particular those which concern income levels. Firstly, evidence suggests that it is advisable to ask questions on annual *disposable* income as this approach is associated with the most accuracy in response (Stopher 2012, p. 180). In addition, the researcher must consider the fact that questions about household income may produce some error because it requires some knowledge on the other individual not participating in the survey. We also note that even when survey participants *know* the answer to a question, they may not want to reveal a truthful answer (Stopher 2012, p. 182).

If the survey is a self-administered survey, and the individual has complete anonymity, this facilitates the elicitation of truthful responses in the context of all types of questions (Stopher 2012, 183). Online surveys offer the perfect environment to preserve respondent anonymity. In presenting our survey to respondents, questions are outlined in multiple choice format. Further, these questions present respondents with a choice of categories (ranges) from which respondents can self-select. Compared to open ended answers, this method is also associated with elicitation of truthful responses, particularly on income levels (Stopher 2012, 186). The data collected on respondents' sociodemographic characteristics is an important part of the empirical analysis. Its accuracy is therefore important.

At estimation stages, interaction terms enable the segmentation of preferences on the basis of observable characteristics. This allows the researcher to identify additional dimensions of differences in preferences between sociodemographic groups. This first section of the survey presents respondents with 10 multiple choice questions on respondents' age group, gender, household income, weekly expenditure on groceries and basic household items, household size, shopping frequency, car ownership, and rate of participation in loyalty schemes. A complete list of the multiple choice questions that were included in the questionnaire can be found below in Table 2.6.

Table 2.6 – List of survey questions presented to respondents on their sociodemographic and household characteristics and shopping preferences

1. How old are you?
   a) 18 – 24
   b) 25 – 44
   c) 45 – 64
   d) 65+

2. What is your gender?
   a) Male
   b) Female

3. How many people live in your household?
   a) 1
   b) 2
   c) 3
   d) 4+

4. What is your primary occupation?
   a) Full-time employment
   b) Part-time employment
   c) Homemaker
   d) Unemployed
   e) Student

5. What is your annual household disposable (after tax) income?
   a) £0 - £13,000
   b) £13,001 - £22,000
   c) £22,001 - £45,000
   d) £45,001 - £65,000
   e) £65,000+

6. On average, how often do you order groceries using the Internet?
   a) Every week
   b) At least once a month
   c) A few times a year
   d) Never

7. On average, how many times per week does your household go grocery shopping?
   a) 1
   b) 2
   c) 3
   d) 4+

8. On average how much does your household spend on groceries and basic household items (toilet roll, bin liners, etc.) each week (excluding alcohol)?
   a) £10 - £35
   b) £36 - £61
   c) £62 - £87
   d) £88 - £103
   e) £104+

9. How many loyalty schemes do you participate in? (Including grocery retailer, cosmetics retailer, airlines etc.)
   a) None
   b) 1-2
   c) 3-4
   d) 5+

10. Do you usually drive a car to go grocery shopping?
   a) Yes
   b) No

2.10 Conclusion

This chapter presented the underlying theory and rationale for using stated choice methods and efficient experimental designs. We presented evidence that orthogonal designs are not optimised for nonlinear discrete choice models, noting that efficiency based designs offer a more compelling alternative. After evaluating the methods available to practitioners, we presented the ways to perform surveys of human populations, noting the possible biases that we would encounter by carrying out the survey online and sampling the population following a convenience sampling approach. Then, we outlined a detailed methodology for the design of a discrete choice experiment to measure consumer preferences for loyalty schemes in the UK groceries market. In doing so we undertook a qualitative assessment of this market and designed an experiment which mimics the features of the sector. We performed a small pilot study to estimate parameter priors to enter the final design of our survey, which is a requirement in *D*-efficient designs. We then also controlled for uncertainty of parameter priors by relying on the Bayesian version of the *D*-error. Chapter III of this thesis outlines and evaluates the method used to collect the data and presents the results of the DCE.

Chapter III
Empirical Results

3.1 Introduction

This chapter presents the empirical results of a discrete choice experiment designed to model heterogeneity in consumer preferences for grocery retailer attributes, namely loyalty schemes. Looking back to the literature review chapter, we found that the economic models used to assess repeat purchase discount type strategies, generally assume that consumers are homogenous in their artificial costs of switching. In other words, when firms implement loyalty schemes, the strategy unilaterally increases the costs of switching for *all* consumers. We found that theoretical models can, and do in many cases, account for heterogeneity in consumer preferences. Typically this is through locational differences in a Hotelling framework. Additional variation may for example, be applied through random variation in preferences across different time periods entering the model. However, until recently, the fact that consumers are likely to be heterogeneous in their switching costs has not been accounted for in the literature (Biglaiser et al. 2016).

In light of this, this thesis set out to re-examine the assumptions entering theoretical models on costs of switching when firms implement loyalty schemes. By using the mixed logit model we are able to test whether the coefficient for the loyalty discount varies between consumers or not. If it does *not*, then it follows that consumers are likely to be homogenous in their artificial costs of switching. This is consistent with the nature of assumptions typically applied in the theory. On the other hand, if consumers' taste for loyalty schemes varies in the population then consumers are likely to be heterogeneous in their artificial costs of switching created by the retailer's strategy. The empirical findings presented in this chapter suggest that loyalty schemes do *not* affect consumers in the same way and that grocery retailers in the UK are likely to compete for consumers across a wide range of price and non-price factors, including loyalty schemes.

Compared to other grocery retailer attributes which exhibit preference heterogeneity, the most variation in households' preferences is for the loyalty scheme discount attribute. Based on our sample of data, around 97% of consumers prefer grocery retailers who offer high levels of customer service in store. However, only around 68% of consumers choose a grocery retailer on the basis of being able to achieve a loyalty discount. This indicates that in the population of grocery shoppers in the UK, only *some* consumers' behaviour is affected by a loyalty scheme when choosing a grocery retailer. This supports the hypothesis we set out in the first chapter which proposes that consumers are likely to incur artificial switching costs *heterogeneously*. This has implications for the theoretical models used to study endogenous switching costs related to loyalty scheme type discounts. For example, our finding implies that loyalty discounts do not unilaterally increase switching costs for all consumers. We also argue that even if a consumer repeatedly buys groceries from the same retailer who offers a loyalty scheme, this will not necessarily mean the individual will then redeem their lump sum coupon. This may be due to personal preferences and consumer seeking to protect their personal data for example. As such, the loyalty scheme may not cost as much to the retailer as implied by some of the models used to study repeat purchase discounts. We also argue that the scheme is unlikely to affect market outcomes in the way suggested by theoretical models. We discuss these aspects of the results at greater length in the discussion section.

The overall structure of this chapter can be summarized as follows. Section 3.2 discusses the benefits and limitations associated with online surveys and the resultant sources of bias we need to consider when undertaking the empirical work. This sub-section also sets out how we propose to test the quality of data collected to assess whether the empirical results are likely to be meaningful. Section 3.3 compares survey respondent characteristics against the UK population statistics along a range of different factors such as location and disposable household income. For example, we find that certain income and age groups as well as UK regions are under or overrepresented in the data. The sampled respondents are proportionality younger, richer and generally more likely to be from more prosperous regions where household incomes sit above the national average. The comparison of sociodemographic characteristics of survey respondents against actual population figures provides the necessary information to determine the relevant frequency weights to be applied to specific groups of consumers found to be underrepresented within our sample. The frequency weights applied to the data are set out at the start of the empirical results section.

Throughout Section 3.4 we present our empirical results obtained by running different model specifications, namely the conditional and mixed logit specifications, with and without interaction terms. The complete list of all the model specifications that are referenced in this chapter can be found in the appendix under table A3.1. Section 3.5 presents additional results to assist in the interpretation of the data, namely, willingness-to-pay estimates for grocery retailer attributes and graphical representations of individual-level parameter estimates. WTP is a useful and alternative way to interpret results by examining ordered preferences in the form of WTP estimates. Individual-level parameter estimates on the other hand, allow the researcher to graphically map how widely (or narrowly) preferences are likely to be dispersed around the mean. Section 3.6 concludes this chapter with a discussion of the relevant insights offered by our findings in the context of the academic literature, implications for competition policy and possible directions for future research.

3.2 Data Collection & Data Quality

This sub-section outlines the process undertaken for the collection of data. In doing so, we set out how we propose to test and control for sources of bias that may impact the quality of our results. In the context of the methodological chapter, we outlined the trade-offs inherent to the different methods available to researchers aiming to collect a representative sample of the population through a survey. We noted in particular the potential sources of bias inherent to data collected via survey methods. These issues, including sources of bias attributed to online surveys, are discussed in greater detail in the methodological chapter Section 2.7. In addition, we also note that Section 2.9 considers how to mitigate bias and elicit truthful responses to sensitive questions such as household income.

The survey outlined in the second chapter of this thesis was uploaded to the online platform Qualtrics and preserved respondent anonymity. The survey itself was made up of two parts, (i) sociodemographic and household characteristics questions; and (ii) the *D*-efficient survey containing 11 hypothetical shopping scenarios. All survey data was collected from the 17th March 2014 up to the 18th July 2017. The purpose of launching the survey online was threefold: (a) elicit truthful responses; (b) adopt a cost and resource efficient method; and (c) facilitate the collection of as large as possible set of survey responses. As the survey was conducted online we benefited from the ability to reach a wide audience. For example, the survey was shared via multiple internal company and university mailing lists as well as social media channels (e.g. Facebook, Twitter and LinkedIn). The survey platform also varied the sequence of questions presented to each respondent to remove sources of bias arising from the ordering of questions. Further, by conducting the survey using Qualtrics enabled us to purchase supplemental responses at very short notice.

Whilst beneficial, our adopted approach was also subject to certain limitations. Firstly, there are specific biases associated with online surveys. Online surveys limit researchers' ability to reach certain respondent groups due internet penetration rates not being 100%. In addition, online surveys require the respondent to have minimum threshold of technological skills. For example, individuals in the older age category are less likely to be savvy internet users are would therefore be less likely to participate in online surveys. On the other hand, certain respondent groups may be overrepresented, for example students and higher income households.

Considering the above, by collecting the data online we were inadvertently either over-including or excluding certain households from participating in the survey. This is consistent with what we observe in the context of the analysis of sociodemographic characteristics of survey respondents presented in the section below. These results show that our sample is significantly under-representative of older age groups and has a large number of students than otherwise suggested by population-level figures. In addition, respondents report very high disposable incomes. We also noted in the methodological chapter the issues associated with asking individuals about household versus personal income. This may increase the scope for reporting errors. These issues are also discussed in the next section of the chapter.

Secondly, we chose to distribute the survey link across as many online channels as possible which was by far the most convenient and least resource intensive way to collect a sufficiently large sample of data for the purposes of our empirical work. In Section 2.6 of the methodological chapter we outlined a number of methods for the collection and sampling of data. In doing so we explained that we would have adopted an alternative approach to the collection of data with access to much bigger resources. This is in terms of availability of time to conduct the research and much greater finances. In this situation, we would have sought to for example, purchase a sampling frame of a representative sample of the UK population and drawn a random sample using one of the random sampling methods available. Alternatively, we could have drawn from a random sample of UK households using random dialling or drawing a random sample from online telephone directory.

We would have then invited individuals for face-to-face interviews by contacting them via post or telephone because face-to-face interviews are said to deliver the most reliable results.[53] Alternatively, on the basis of a random sample of grocery stores we could have interviewed a random sample of people leaving/ entering each selected grocery store. This would have also required face-to-face interview techniques.

[53] These issues are discussed at greater length in Section 2.6 of the methodological chapter.

However, as we explained in greater detail in Section 2.6, there are a number of difficulties associated in drawing a completely random and representative sample of households on the basis of postal addresses in a telephone directory or other comparable methods. This approach also represents the most resource (time and money) intensive method for collecting the survey data, which was one of the main considerations when choosing between different survey methods (Stopher 2012, p. 362).

Our chosen approach, an anonymous online survey, is now a common method used to collect survey data. Online surveys are known to be more effective in eliciting truthful responses by ensuring respondents are able to enjoy complete anonymity when submitting their responses. They also permit the careful design of questions to further facilitate the elicitation of truthful responses. This was particularly relevant as our survey included more sensitive questions in relation to income. We note that extracting truthful information on income levels from respondents in surveys is a well-known issue in survey methods. However, we were able to minimize this risk by conducing the survey online and presenting available answer as ranges in multiple choice format. As noted above, Section 2.6 discusses this issue in greater detail.

The collection of responses was carefully monitored throughout the data collection process. In the first instance, the link to the survey was shared via a number of online channels in the UK. Furthermore, additional responses were purchased through the Qualtrics panel of respondents. The initial data gathering exercise enabled us to collect 293 responses. Through the Qualtrics panel we purchased a further 142 responses which were recorded during the same time frame as the first sample, adding up to a total of collected 435 responses. As explained in the methodological chapter, researchers generally aim to achieve a sample size of 200 or 300 to achieve robust parameter estimates in the context of discrete choice experiments (Rose and Bliemer 2013, Stopher 2012, p. 256).

After downloading the data from Qualtrics in .csv format we performed a number of checks on the data for quality assurance purposes. Of the total 435 recorded responses, 8 responses were removed from the total sample for one of three different reasons. Firstly, we removed all respondents who completed the survey in under 5 minutes. This was to exclude individuals not actually reading the instructions or questions presented to them in the survey. Secondly, we removed a survey respondent who selected the same option during each survey question. We also determined the approximate locations of survey respondents at the time of taking the survey using their

GeoIP location. In doing so we identified and removed the survey responses from Northern Ireland where the grocery retail sector is significantly different to England, Scotland and Wales.[54] This left us with a total of 292 and 135 remaining responses from the first and second samples respectively.

After performing these initial quality controls we sought to verify whether we could identify and control for sources of bias in our data which would likely affect the quality of our results. In doing so we noted that a good quality and unbiased sample of the population would be representative of UK household characteristics and consistent with actual UK household preferences. This would enable us to estimate the data to obtain reasonable empirical results with reliable real-world applications. Considering the above we performed a series of checks to test the quality of the data, specifically:

(i) compared the implied market shares computed on the basis of survey responses to actual grocery retailer market shares in the same time period as when the survey was conducted;

(ii) compared the survey respondents' household and sociodemographic characteristics to publically available statistics on the UK population with a particular focus on known drivers of grocery shopping preferences including but not limited to their household income, age and household size (*see Section 3.3*);

(iii) compared the first and second group of respondents with reference to their household and sociodemographic characteristics to test for material differences between these two samples (*see Section 3.3*); and

(iv) on the basis of the comparative assessment in (ii) apply population based weights to the data to address potential sources of bias *(see Section 3.4)*.

To address bullet (i) above, we computed market shares by taking the frequency of responses attributed to retailers A, B, C and D as a proportion of total responses and compared these against actual grocery retailer market shares. Even though the survey was "unlabelled", in the sense that respondents did not choose between identifiable

[54] See for example a recent report by Kantar Worldpanel on the Northern Ireland Retail Landscape which sets out the key players in the market and their corresponding market shares. The market structure outlined in this report is very different to the UK grocery retail landscape. For example, Tesco has nearly 35% market share (in the UK <30%) and retailers Morrison's and Waitrose do not appear to operate in Northern Ireland. https://www.food.gov.uk/sites/default/files/coracampbellkantarworldpanel.pdf, Accessed 9th October 2017.

retailer names, the experiment was designed on the basis of actual retailer offerings. In this context, retailer profiles A, B, C and D were modelled against actual grocery retailer features of Tesco, Sainsbury's, Asda and Waitrose respectively. Considering the above, the similarity between the survey and actual market shares is a good indicator of the quality and representativeness of the data. The market shares computed on the basis of the chosen alternatives are presented below in Table 3.1. We also include a so-called "group of four market share". In other words, market shares computed as a proportion of the total 67.5% share attributed to these four retailers in the UK groceries market.

Table 3.1 – Implied and Actual Grocery Retailer Market Shares in 2014

	Tesco	**Sainsbury's**	**Asda**	**Waitrose**	**Total**
Market Share (sample)	45.1%	17.5%	28.4%	9.0%	100.0%
2014 Market Share (actual)	28.8%	16.2%	17.4%	5.1%	67.5%
2014 Group of Four Market Share	42.7%	24.0%	25.8%	7.6%	100.0%

Notes: Survey market share figures were computed by calculating the proportion of times retailers labelled as A, B, C or D were selected by respondents during the survey. Actual market share figures are sourced from Kantar. [55] Group of four market shares were computed by assuming the grocery market was restricted to only the four retailers listed above taken as a proportion of actual market shares in 2014.

The market shares in the first row of Table 3.1 above, are consistent with actual market shares observed in real world markets both in terms of the overall split and relative sizes of retailers. Firstly, comparing retailers' relative sizes to each other, in terms of market shares, Tesco, Asda, Sainsbury's and Waitrose are the first, second, third and fourth largest retailers in order of market share magnitudes respectively. This is the case when looking at both the sample and the actual UK market. In addition, there is overall limited variation between the sample market shares and the actual market shares. For example, the actual market shares of Tesco, Asda and Waitrose are less than 3% higher compared to the survey based market shares. On the other hand, Sainsbury's market share is 6.5% lower in the sample than its actual market share. In conclusion, although the implied survey market shares are not strictly identical to what we observe in actual markets, they are very similar. On balance, the results of this check are consistent with what would be expected of good quality data that is

[55] http://uk.kantar.com/consumer/shoppers/2014/2309-kantar-worldpanel-uk-grocery-share-data-september/

representative of actual UK household preferences and therefore more likely to deliver reasonable empirical results.

The next section evaluates respondent characteristics against those of the UK population. This data is also disaggregated by respondent group to assess whether any significant differences exist between these two sets of respondents. In doing so, we are able to perform quality checks (ii) and (iii) described above. The results of this evaluation indicate that although there is some degree of variation between the two sets of respondents, on balance the differences are not substantial. Instead, more material differences arise between certain sociodemographic characteristics of survey respondents and those found in UK population. These aspects of the data and how we propose to control for these sources of bias are discussed at greater length below.

3.3 Demographic Characteristics of Survey Respondents

In section 2.10 above we outlined the sociodemographic questions that were included in the final survey design. The inclusion of these specific questions enabled us to collect respondent-level data which can be segmented according to respondents household and sociodemographic characteristics. Ultimately, the segmentation of respondents allows us test for the presence of preference heterogeneity between groups of consumers/ households through interaction terms contained in the various model specifications.[56] In addition, this segmentation allows us to evaluate the quality of the data by comparing the characteristics of the two groups of survey respondents to each other and also to those of the general UK population. In doing so, we test for sample representativeness and sources of likely bias that may be corrected through the application of population weights. For example, as noted above the survey was conducted online rather than using the traditional methods like telephone interviews or via postal delivery. This approach to the collection of data may have led to the overrepresentation of younger and more educated individuals and the underrepresentation of less technologically sophisticated shoppers and also older individuals. We address this type of issue further below through the comparative assessment.

A key assumption we make in interpreting our results is that the choices made by individual respondents also account for the preferences of the household they live in. We note that individuals in the population are likely to either:

[56] Please see the appendix for a full list of model specifications and interaction term descriptions.

- be the main person responsible for their household's shopping; or
- may share grocery shopping responsibilities with others in the household; or
- may generally shop as a household; or
- may not go grocery shopping at all.

Being mindful of the above, as part of the preliminary survey questions, survey respondents were asked whether they were responsible for performing the main shop on behalf of their household. The survey instructions also clearly set out that the questionnaire related to household behaviour. Additionally, the remaining questions in the survey were also formulated to ensure the respondent would consider their overall household spending and shopping habits. Thus we interpret the estimated coefficients as representing the preferences of individual shoppers (consumers) as well as the households they live in. We also note that some individuals live by themselves in which case this point is not pertinent. However as we show below, a majority of survey participants do not live by themselves.

Tables 3.2-3.6 below contain a number of statistics on respondent sociodemographic characteristics and shopping behaviour. We compare these figures to UK population statistics which have been sourced from a number of organisations such as the Office for National Statistics, HM Revenue & Customs, The Institute of Grocery Distribution ("IGD") and Kantar Worldpanel ("Kantar").[57] Most statistics, unless otherwise stated, are representative of the year 2014 which corresponds to the year the survey was conducted.

[57] A complete list of data sources can be found in the bibliography.

Table 3.2 – Survey Summary Statistics: General Household Characteristics

Demographic Variable	Description	Respondent Group 1 (%)	Respondent Group 2 (%)	All Respondents (%)	UK Population (%)
Age	18 – 24	16.44	8.89	14.05	11.54
	25 – 44	61.30	58.52	60.42	33.63
	45 – 64	19.18	25.93	21.31	32.37
	65+	3.08	6.67	4.22	22.47
Gender	Male	41.78	36.30	40.05	49.22
	Female	58.22	63.70	59.95	50.78
Household Size	1	17.12	11.85	15.46	27
	2	39.04	24.44	34.43	36
	3	19.86	31.11	23.42	17
	4+	23.97	32.59	26.70	20.00
Primary Occupation	Full-time employed	68.84	47.41	62.06	60.01
	Part-time employed	8.22	23.70	13.11	22.04
	Homemaker	1.71	14.81	5.85	6.27
	Unemployed	1.37	10.37	4.22	5.16
	Student	19.86	3.70	14.75	6.52
Disposable Household Income	£0 - £13,000	11.64	14.07	12.41	17
	£13,001 - £22,000	14.73	23.70	17.56	40
	£22,001 - £45,000	40.75	45.19	42.15	35
	£45,001 - £65,000	15.41	10.37	13.82	4
	£65,000+	17.47	6.67	14.05	4

Notes: These figures are based on 427 survey responses. Respondent groups 1 and 2 have 292 and 135 survey responses respectively. Figures may not add up to 100 due to rounding. Data for age, gender, household size and occupation was sourced from the ONS. Data on disposable household income was sourced from HM Revenue & Customers. All data relates to 2014.

Let us first consider the distribution of respondents across the various age categories. Both respondent groups are similar to each other, however, there are important differences between the distributions of age groups in the sample groups compared to the UK population. In particular the 65+ age category is much smaller than other age categories in our sample and also in terms of what is observed within the UK population. Instead, our sample contains nearly twice as many individuals in the 25-44 age category than suggested by population-level data. Furthermore, our sample contains relatively more female respondent than male respondents. This may be due to the fact that the survey was intended to be completed by the primary shopper of that household and women are more likely to carry out this role. For example, consider the findings of a large study conducted by the Food Standards Agency published in 2007. "*Consistently throughout the period of the study, more women (77% in 2006) took all/most of the responsibility for household food shopping compared to men (29% in 2006).*" [58]

The population statistics for both age and gender were obtained using the ONS's Population Analysis Tool 2014. Employment statistics were derived from the 2013 Labour Force Survey ("LFS") published by the ONS. This data set contains information on different types of labour market activity/ inactivity for individuals aged 16-64. We chose the LFS dataset as it provided a breakdown of labour market activity by type. However, the LFS dataset allowed for multiple occupations. So for example, certain individuals were classified both as part-time employed and as students. On the other hand, our questionnaire did not allow for respondents to select more than one answer per question. Furthermore this data set contained statistics for individuals aged 16-64, while our survey was completed by individuals aged 18 to 65+. Considering the above, we are unable to make a like for like comparison to for this specific category and we therefore do not place significant weight on the underlying reasons differences observed between our sample and the population level statistics.

Let us first consider the differences between the two samples of respondents in the occupation category. The primary occupations of survey respondents are visibly different between the two groups 1 and 2. A number of the differences in primary occupation between the two sets of respondents, could be driven by the different ways the two samples were collected. The first sample was collected by sharing the survey via mailing lists and social media. On the other hand, the second set of responses were collected via a paid service which required paying a panel of respondents to complete the survey. For example, the second sample contained around twice as many unemployed individuals compared to what we observe in the national statistics. This is consistent with the fact that someone who is employed in a full-time job would be less likely to participate in unconventional side-jobs like being a part-time panellist for online surveys. Furthermore, there are a large number of student respondents in the first respondent group which is consistent with the channels that were used to distribute

[58] We note that this study is ten years old and that more men are likely to be performing the household shopping in 2017.
http://webarchive.nationalarchives.gov.uk/20111206144033/http://www.food.gov.uk/multimedia/pdfs/cas07uk.pdf

the survey, namely a number of university mailing lists. When taking the respondent groups together, on balance the combined figures are representative overall in the context of full-time employment and unemployment. These figures are very similar to the UK population statistics in the fourth column.

In terms of household sizes, the data we collected is under representative of single person households. However, it is worth noting that 37% of respondents who participated in the survey were located in London where living costs are higher than the rest of the country. As a consequence of higher living expenses and well-known supply side shortages, more individuals are likely to live in shared accommodation (i.e. as part of a larger household) in London which could be driving the above statistic. The locations of survey respondents may also be driving the way household income categories are distributed in the sample, in particular the first sample group.

Over 17% of respondents in the first group stated their household earnings after tax were over £65k whereas this figure was 7% and 4% for the second group and UK population respectively. Table 3.5 further below presents the gross disposable income by UK region showing that London and the East of England have the highest average disposable household incomes compared to other regions. In addition, Table 3.4 contains locations of respondents at the time of completing the survey indicating that the majority of our respondents were also from these higher income regions. The reported household incomes are higher than the national average and there are at least two biases driving this result. The first is the risk that respondents will be uncertain about their household income as this information is about another individual. Secondly, another source of bias is the fact that the survey was carried out online which restricts the sampling to more affluent individuals as discussed in more detail in the methodological chapter.

Table 3.3 – Survey Summary Statistics: Household Shopping Behaviour

Grocery Shopping Variable	Description	Respondent Group 1 (%)	Respondent Group 2 (%)	All Respondents (%)	UK Population (%)
Shopping Frequency (weekly)	1	37.33	31.85	35.60	-
	2	29.79	41.48	33.49	-
	3	18.49	14.81	17.33	-
	4+	14.38	11.85	13.58	-
Online Shopping Frequency	every week	7.53	15.56	10.07	2.43
	few times a month	13.36	31.85	19.20	11.93
	a few times a year	29.79	22.96	27.63	7.74
	Never	49.32	29.63	43.09	77.9
Weekly Expenditure (food & non-food nondurable household items)	£10 - £35	17.81	14.81	16.86	-
	£36 - £61	28.08	28.15	28.10	-
	£62 - £87	25.00	25.93	25.29	-
	£88 - £103	15.75	20.74	17.33	-
	£104+	13.36	10.37	12.41	-
Loyalty Scheme Participation (all types)	0	17.12	5.19	13.35	-
	1 – 2	49.66	46.67	48.71	-
	3 – 4	22.26	33.33	25.76	-
	5+	10.96	14.81	12.18	-
Drives a Car to Buy Groceries	Yes	54.45	67.41	58.55	61
	No	45.55	32.59	41.45	39

Notes: These figures are based on 427 survey responses. Respondent groups 1 and 2 have 292 and 135 survey responses respectively. Figures may not add up to 100 due to rounding. Sources for the population-level figures can be found in the data sources section of the bibliography.

Table 3.3 above presents the summary statistics for grocery shopping related questions, such as loyalty scheme participation and shopping frequency. There are a number of differences between the two groups of respondents throughout these different categories. Let us consider them in turn. Firstly, the second group of respondents are more avid loyalty scheme participants than the first group. There are in fact, 17% of respondents in the first group who stated they did not participate in any loyalty scheme at all. The corresponding figure for the second group was 5%. In addition, there are significantly more individuals in the second group of respondents who stated that they typically drive a car to go grocery shopping compared to those in the first group. The combined figures however, are in line with the population level statistics.

Looking at the distribution of responses in relation to preferences for online shopping, the sampled respondents appear to purchase groceries online more frequently than the national average. This holds in particular for the second group of respondents. If we compare the overall statistics for online shopping preferences to figures published by Kantar, proportionately twice as many of our sampled respondents regularly purchase groceries online. This could be due to individuals in our sample working demanding jobs with less time to shop or simply they may prefer alternative shopping channels. This could also be attributed to the fact that the sample contains a larger number of younger individuals than the national average and the survey itself was conducted online which, as noted above, may attract a specific type of respondent.

On the basis of the table above, once-stop-shopping remains the preferred shopping method among roughly a third of survey respondents. Another third of respondents state that they go grocery shopping at least twice per week. The remaining 30% go grocery shopping at least 3 times per week. Although we could not find like-for-like data on grocery shopping frequency, we refer to an IGD report which looks at developments in grocery shopping frequency over time for some insights. The IGD shopper insight report focuses on the growth in 'top-up shopping' through the analysis of survey data on consumer shopping behaviour.[59] In the report, top-up shopping is defined as a smaller shopping trip in terms of basket size, which is carried out by customers wishing to top-up their main grocery shop. According to IGD, 46% of respondents in the survey claimed that they were top-up shopping more often that 2-3 years ago. The report highlights that consumers used to be more likely to carry out one major shop per week and preferred grocery retailers offering a so-called one-stop-shop for all major items they needed for that week. In turn, the groceries sector experienced growth of big out of town supermarkets. More recently, the sector has seen growth of convenience stores in line with consumers favouring more frequent shopping trips over a one-stop-shop. This is consistent with what we observe in our data.

[59] IGD ShopperVista Report, 'Top Up Shopping', June 2015

Table 3.4 – Locations of survey respondents

Location	Respondent Group 1 (%)	Respondent Group 2 (%)	All Respondents (%)	UK Population (%)
East of England	39.38	2.22	27.63	9.59
East Midlands	1.03	3.70	1.87	7.39
London	40.75	27.41	36.53	13.61
North East	1.03	4.44	2.11	4.17
North West	5.48	13.33	7.96	11.37
Scotland	1.71	6.67	3.28	8.52
South East	5.14	9.63	6.56	14.14
South West	1.71	10.37	4.45	8.64
Wales	0.34	5.19	1.87	4.93
West Midlands	1.71	11.11	4.68	9.10
Yorkshire & the Humber	1.71	5.93	3.04	8.54

Notes: Locations of 427 survey respondents were recorded at the time of survey completion. Respondent groups 1 and 2 have 292 and 135 survey responses respectively. Figures may not add up to 100 due to rounding. Population-level figures exclude N. Ireland and have been sourced from the ONS.

Using the ONS's population analysis tool, we calculated the proportion of individuals living in different UK regions. This data is presented in Table 3.4 above. The figures which represent the second group of respondents are significantly more representative of the UK population statistics than those attributed to the first group. In the first group, there are more individuals located in London and the East of England compared to the UK average. In fact there are more than twice as many individuals from London and the East of England in our sample compared to the UK demographic. We noted above, that this figure is also related to the high income levels reported by survey respondents considering that regional differences are an important determinant of income of households. Table 3.4 below presents the respective regional gross disposable household incomes ("GDHI") by region as well as respondent income group by respondent location. The GDHI regional averages presented in the above table represent average incomes after taxation and other social contributions have been deducted at the individual-level *not* the average of household incomes in different regions.

Table 3.5 –Gross disposable household income (“GDHI”) by UK region and survey respondent income group by respondents’ location

Region	G.D.H.I. (£)	£0 -£13000 (%)	£13,001-£22,000 (%)	£22,001-£45,000 (%)	£45,001-£65,000 (%)	£65,000+ (%)
United Kingdom	17,965	-	-	-	-	-
East of England	18,897	28.30	28.00	30.00	30.51	16.67
East Midlands	16,217	1.89	2.67	1.67	1.69	1.67
London	23,607	39.62	29.33	33.89	32.20	55.00
North East	15,189	3.77	2.67	2.22	1.69	0.00
North West	15,776	7.55	13.33	6.11	11.86	3.33
Scotland	17,095	0.00	1.33	3.89	6.78	3.33
South East	20,434	3.77	5.33	7.78	5.08	8.33
South West	18,144	3.77	5.33	3.89	6.78	3.33
Wales	15,302	1.89	4.00	1.11	1.69	1.67
West Midlands	15,611	1.89	6.67	6.11	0.00	5.00
Yorkshire & the Humber	15,498	7.55	1.33	3.33	1.69	1.67
Total	-	*100%*	*100%*	*100%*	*100%*	*100%*

Notes: figures are for 2014 and have been sourced from the ONS. Locations of 427 survey respondents were recorded at the time of survey completion.

We had initially suspected that the high incomes reported by respondents were likely driven by regional differences. Considering the statistics presented in the table above, over half of respondents who reported household incomes of £65k+ were located in London. On the other hand, 40% of individuals in the lowest income category were also located in London. The second largest group of individuals in the highest income category were located in the East of England. This is among the more prosperous regions on the basis of the GDHI figures. Considering another example, the North East is reported as having the lowest average household incomes in the UK. None of these survey respondents located in the North East reported as being in the highest household income category. The collected data therefore captures both the income inequalities prevalent in the capital as well as regional household income disparities across UK regions captured in the first column above. This finding is consistent with good quality of data which is likely to be more representative of preferences within the UK population.

Table 3.6 –Average weekly household expenditure on food and non-alcoholic drink by UK region and survey respondent average spending by respondents' location

Region	Household average weekly spending (£)	£10 - £35 (%)	£36 - £61 (%)	£62 - 87 (%)	£88 - 103 (%)	£104+ (%)
United Kingdom	58.20	-	-	-	-	-
East of England	61.70	26.39%	28.33%	27.78%	25.68%	30.19%
East Midlands	57.80	1.39%	0.83%	1.85%	1.35%	5.66%
London	62.60	45.83%	34.17%	35.19%	32.43%	37.74%
North East	49.60	1.39%	0.83%	2.78%	4.05%	1.89%
North West	54.80	6.94%	8.33%	8.33%	8.11%	7.55%
Scotland	56.10	2.78%	0.83%	3.70%	5.41%	5.66%
South East	64.20	1.39%	7.50%	8.33%	10.81%	1.89%
South West	60.40	2.78%	6.67%	3.70%	4.05%	3.77%
Wales	53.20	1.39%	4.17%	0.00%	1.35%	1.89%
West Midlands	55.80	6.94%	3.33%	6.48%	2.70%	3.77%
Yorkshire & the Humber	51.40	2.78%	5.00%	1.85%	4.05%	0.00%
Total	-	*100%*	*100%*	*100%*	*100%*	*100%*

Notes: figures are for 2014 and have been sourced from the ONS. Locations of 427 survey respondents were recorded at the time of survey completion.

Table 3.6 above presents the average weekly household expenditure on food and non-alcoholic drinks by region. The table also includes average weekly household expenditure reported by survey respondents by respondent location. As above for the comparison of regional incomes, the levels of spending on groceries and non-alcoholic drink reported in the survey are consistent with the levels of expenditure on these items in different regions. This is also consistent with average household expenditure on groceries varying between households enjoying different levels of disposable income. The next section considers the empirical results we obtained when fitting the data to a number of discrete choice model specifications. In doing so, we also explain the weights applied to the data on the basis of population statistics presented in the various tables above.

3.4 Fitting the Data to the Conditional and Mixed Logit Models

Logistic regression models are frequently used by researchers to estimate data obtained through discrete choice experiments. As outlined in Section 2.3 of the methodological chapter, this family of models allows the researcher to assess consumer decision-making and preferences in a target population. Developed over 30 years ago, McFadden's (1974) conditional logit model has been the most widely used across disciplines. The extension of conditional logit model to the mixed logit specification represents a more recent development in discrete choice modelling.[60] The results presented in this section were estimated by fitting the survey data to the conditional and mixed logit models in *Stata* using the commands *clogit* and *mixlogit* respectively.

The main difference between the two model forms is that the conditional logit assumes individuals in the population share the same preferences. In other words, this implies that the computed mean coefficients for model attributes are fixed regardless of the individual. However, when estimating the data with the conditional logit specification, researchers can include interactions (i.e. covariates) between explanatory variables and sociodemographic variables to test for preference heterogeneity (Hole 2008). We chose to test 39 interactions terms in total in the context of the empirical work for both the conditional and mixed logit models. These results are presented further below.

The interaction terms were chosen on the basis that there may be some variation in preferences between certain sociodemographic groups. Specifically, we are interested in understanding whether differences in gender, household composition, occupation, age, shopping frequency, loyalty card ownership and household income account for variation in preferences for grocery retailer attributes. In addition to being able to capture preference heterogeneity, we also rely on the analysis with interaction terms to assess whether our results are reasonable and reliable in the sense that they are consistent with general underpinnings of microeconomic theory and general common sense. For example, on this basis, we anticipate that the interactions will capture differences in household income and/ or household size which are typically important drivers of spending habits.

Compared to the structure of the survey questions, for the purposes of the empirical work we have redefined new categories of sociodemographic groups in order to preserve degrees of freedom. For example, in the context of the survey, respondents were asked whether their ages corresponded to a set range in one of four age categories. Whereas, for the purposes of the empirical analysis, we restrict age to two categories only: individuals aged 18-45 ("younger") and individuals aged 45 and above ("older").

[60] Refer to Section 2.3 of the methodological chapter for more details on McFadden's conditional logit model and extension to the mixed logit model specification.

Similar reconfigurations have been performed to design the other interaction terms and a complete list of these interactions and their corresponding descriptions are all outlined in the Appendix under Table A.3.2.

The main benefit of the mixed logit model in this specific context, is that preference heterogeneity can be captured not only through sociodemographic covariates but also through the magnitudes of the estimated standard deviations of the random coefficients themselves. The magnitudes of the standard deviations computed by running a mixed logit specification, indicate how preferences are distributed among individuals in the population of interest. For example, further below in the next section of this chapter we include individual-level parameter kernel density graphs depicting the likely distribution of preferences (wide vs narrow) around the mean coefficient computed through a simulation of responses on the basis of individual-level preferences.

Ahead of fitting our data to the conditional and mixed logit model, we sought to control for likely sources of bias through the application of weights. We impose weights on the age, income and location categories computed on the basis of population-level statistics outlined in Table 3.7 below.[61] For the purposes of our empirical analysis we sought to collect a representative dataset of UK households across the various sociodemographic categories. In particular, to obtain meaningful results, we wanted to ensure that the data would be representative of the population in terms of known drivers of preferences and spending habits in the population. In addition, given the UK dimension of our study, it was also important for regional differences in the sample to be consistent with what is observed in population-level data. In the previous section we compared the sociodemographic characteristics of survey respondents to those characteristics found in the general population. In doing so, we found that certain categories of households and regions were over or underrepresented in our sample. We note that this may not necessarily be an issue for certain types of studies. However, when studying preferences which exist within the general population, or which can be attributed to a specific type of consumer, the quality

[61] Both Stata commands *clogit* and *mixlogit* accommodate weights through the sub-command *fweight.* See the Stata manual for more details for the description of the different types of weights accommodated by the software, https://www.stata.com/manuals13/u11.pdf#u11.1.6weight, (accessed 24/10/2017)

of the study depends on whether the sample is representative of the population of interest.

When population-level characteristics are known, it is possible to compute frequency weights to be applied to specific sociodemographic variables in the sample of data.[62] In other words, frequency weights replicate choices made by individual respondents associated with those characteristics which need to be adjusted through weighting. The researcher must compute the ratio of the survey to population values to determine the appropriate weights to be applied (Stopher 2012, p. 426). The previous section presented the population level statistics which have been used to compute the relevant population weights to be applied to the data. We also note that the *Stata* commands *clogit* and *mixlogit* both accommodate frequency weights through the *fweight* sub-command. Table 3.7 below presents the frequency weights determined on the basis of the ratio between the survey and population statistics.

Table 3.7 – Population Based Frequency Weights

Demographic Variable	Survey (%)	Population (%)	Frequency Weight
Age group: Aged 65+	4.22	22.47	5
Income group: £13,001-£22,000	17.56	40	2
Location: North East	2.11	4.17	2
Location: East Midlands	1.87	7.39	4
Location: Scotland	3.28	8.52	2
Location: South East	6.56	14.14	2
Location: South West	4.45	8.64	2
Location: Wales	1.87	4.93	2
Location: West Midlands	4.68	9.10	2
Location: Yorkshire & the Humber	3.04	8.54	3

The results presented in this section were derived in *Stata* after applying the weights outlined in the table above. However, for comparative purposes, we include some specifications in the Appendix which do not apply population weights namely models CL1 (b) and ML2 (b). A list of all the model specifications and descriptions

[62] Stopher, P., Collecting, Managing, and Assessing Data Using Sample Surveys, Chapter 19 "*Data expansion and weighting*", Cambridge University Press, 2012.

discussed in this section can be found in the Appendix under table A.3.1. In addition to the application of weights, when running the various specifications in *Stata* we instruct the software to cluster standard errors at the individual-level.[63] In other words, this restriction imposes the assumption that a single individual's preferences do not vary between choice situations, however, they can vary *between* individuals. This restriction ensures that coefficient estimates produce robust standard errors. We note that this does not affect the magnitudes of the estimated coefficients.

The survey data was set-up to contain an individual respondent identifier code captured by variable *pid*. Therefore when running the conditional logit model we are able to apply the sub-command *vce(cluster pid)* to cluster standard errors at the individual-level. Similarly for the *mixlogit* command we use the sub-command *cluster(pid)* to cluster standard errors at the individual-level. Without clustering, the corresponding standard errors and *z*-statistics are artificially too low and too high respectively compared to their true values.[64] As above, for comparative purposes some of the results presented in the Appendix do not control for robust standard errors. However, we note that all the results presented in the main body of this chapter contain both robust standard errors and apply population weights.

This sub-section is structured as follows. We begin by presenting the results obtained when fitting the data to the conditional logit model specification, with and without interaction terms. We then build on the complexity of these models by fitting the data to a number of different mixed logit specifications. For example, we test the impact of imposing different distributional assumptions on certain variables, including the log-normal and standard normal distributions. In evaluating the results, we compare the goodness of fit of the different models by performing the likelihood ratio test using the χ^2 distribution.

Among the results presented in this section, we place the most emphasis on the estimates obtained when fitting the data to two mixed logit model specifications ML2 and ML4. The first is the mixed logit model without interaction terms and the second includes statistically significant interaction terms. However, as noted above, we begin

[63] See guidance on clustering standard errors in Stata in different contexts contained in the following presentation "Clustered Errors in Stata" https://www.stata.com/meeting/13uk/nichols_crse.pdf, (accessed 24/10/2017)

[64] See for example the differences in estimated standard errors between CL and CL (a) where model CL clusters standard errors at the individual-level and model CL (a) does not. Standard errors are higher for model CL than CL (a).

this section by considering the less complex model specifications. As such, we first consider the results in Table 3.8 for conditional logit models CL1 without interactions and CL2 with interactions between the explanatory variables and respondent sociodemographic characteristics.

Table 3.8 – Conditional logit models CL1 & CL2

	CL1			CL2		
Variable	**Coef.**	**|z-stat|**	**O.R.**	**Coef.**	**|z-stat|**	**O.R.**
Price	-0.054*** (0.006)	9.34	0.947	-0.046*** (0.015)	3.05	0.955
Discount	0.009*** (0.001)	6.17	1.009	0.017*** (0.005)	3.73	1.017
Time	-0.061*** (0.009)	6.91	0.941	-0.111*** (0.023)	4.79	0.895
Medium Quality	0.558*** (0.122)	4.58	1.746	0.562*** (0.128)	4.39	1.754
High Quality	0.555*** (0.138)	4.02	1.741	0.558*** (0.148)	3.77	1.747
Very High Quality	0.830*** (0.150)	5.54	2.292	1.015*** (0.202)	5.04	2.760
Medium Range	0.707*** (0.105)	6.72	2.027	0.898*** (0.115)	7.78	1.978
High Range	0.924*** (0.127)	7.28	2.519	0.833*** (0.145)	5.75	2.454
Very High Range	1.047*** (0.112)	9.38	2.850	1.199*** (0.154)	7.8	2.588
Medium Service	0.845*** (0.107)	7.90	2.327	0.682*** (0.106)	6.45	2.455
High Service	0.988*** (0.122)	8.11	2.687	0.898*** (0.129)	6.97	2.300
Very High Service	1.152*** (0.144)	7.99	3.165	0.951*** (0.209)	4.55	3.316
Female*Price				-0.012 (0.011)	1.11	0.988
Female*Discount				0.002 (0.003)	0.68	1.002
Female*Time				-0.024 (0.019)	1.27	0.976
Female*VH Quality				-0.474** (0.192)	2.46	0.623
Female*VH Range				0.013 (0.168)	0.08	1.013
Female*High Service				0.077 (0.166)	0.47	1.080
Large Household*Price				0.017 (0.011)	1.52	1.017
Large Household*Discount				0.005* (0.003)	1.89	1.005
Large Household*VH Range				-0.129 (0.169)	0.76	0.879
Unemployed*Price				-0.019 (0.014)	1.31	0.981
Student*Price				0.013 (0.012)	1.15	1.013
Unemployed*Discount				0.002 (0.005)	0.31	1.002
Student*Discount				0.0004 (0.004)	0.1	1.000
Unemployed*Time				0.0158 (0.035)	0.45	1.016

Student*Time		-0.038	1.49	0.963
		(0.025)		
No Car*Time		-0.0055	0.31	0.995
		(0.017)		
No Car*VH Range		0.281	1.55	1.324
		(0.181)		
18 – 44 Age Group*Price		-0.026**	2.04	0.974
		(0.013)		
18 – 44 Age Group*Discount		-0.001	0.48	0.999
		(0.003)		
< £22,000 HI*Price		-0.016	1.41	0.984
		(0.011)		
> £45,000 HI*Price		0.021	1.33	1.021
		(0.016)		
> £45,000 HI*Discount		-0.004	0.94	0.996
		(0.004)		
> £45,000 HI*VH Quality		0.356	1.52	1.427
		(0.234)		
> £45,000 HI*High Service		0.542	3.07	1.719
		(0.176)		
< £22,000 HI*Discount		-0.003	1.08	0.997
		(0.003)		
> £45,000 HI*Time		0.036	1.58	1.037
		(0.023)		
Frequent Online Shop*Price		0.044***	4.08	1.045
		(0.011)		
Frequent Online Shop*Discount		-0.001	0.22	0.999
		(0.003)		
Frequent Online Shop*Time		0.070***	4.00	1.073
		(0.018)		
Infrequent Online Shop*Price		-0.010	0.89	0.990
		(0.011)		
Infrequent Online Shop*Discount		-0.003	0.99	0.997
		(0.003)		
Infrequent Online Shop*Time		0.026	1.41	1.027
		(0.019)		
Infrequent Online Shop*VH Range		-0.014	0.08	0.986
		(0.180)		
No Loyalty Cards*Price		0.008	0.66	1.008
		(0.013)		
No Loyalty Cards*Discount		-0.014***	3.52	0.987
		(0.004)		
No Loyalty Cards*Time		0.040*	1.71	1.040
		(0.023)		
1-2 Loyalty Cards*Time		0.020	1.05	1.020
		(0.019)		
1-2 Loyalty Cards*Discount		-0.005	1.64	0.995
		(0.003)		
1-2 Loyalty Cards*Price		-0.003	0.21	0.997
		(0.012)		
Log-likelihood	-5204.184	-4942.885		
Nr. of Resp.	427	427		
Nr. of Obs.	18,832	18,832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

The above table shows that the estimates for model CL1 are all statistically significant at the 1% level. The estimates show that prices are an important determinant of grocery retailer choice, however the non-price store characteristics also impact the odds of a consumer choosing a given retailer. Let take a closer look at these results. The above table contains the mean coefficients of the explanatory variables with the robust standard errors presented below these values in parentheses. The table also includes columns for *z*-statistics and the mean coefficient values transformed to odds ratios. The *z*-values which correspond to the mean coefficients indicate the relative explanatory power of the various attributes in respondents' choice of grocery retailer. The attributes with the largest *z*-values are the grocery basket price, very high level of product range, high level of service, loyalty scheme discount and time travelling to the store. The relative magnitudes of the *z*-statistics in model CL1 indicate that presence of a very high range of products in a grocery store and the average basket price account for the most variation in choice of grocery retailer.

The mean coefficient estimates themselves for models CL1 and CL2 in Table 3.8 are in the form of log of odds ratios. These values can be interpreted as relative weights that are assigned to the attributes in the underlying utility function set up by the analyst. At times researchers refer to the coefficient estimates as probabilities but these are in fact *logit* probabilities and should be interpreted accordingly. A basic logistic regression model can be expressed as $\boldsymbol{logit}(\boldsymbol{p})=\boldsymbol{log}(\boldsymbol{p}/(1-\boldsymbol{p})) = \boldsymbol{\beta}_0\mathbf{x}_0 +\ldots+\boldsymbol{\beta}_k\mathbf{x}_k$, where $\boldsymbol{p}$ is the overall probability of choosing a given retailer and $\boldsymbol{p}/(1-\boldsymbol{p})$ expresses this probability in the form of an odds ratio.[65] To facilitate the interpretation the estimated coefficients we have calculated the odds ratios by computing the exponent of the mean coefficient values using the post-estimation "*or*" command in *Stata 12*. The ***log***(odds) is converted to odds by taking the exponent of the coefficients produced by Stata for attribute *k*, in other words calculating e^{β_k}. We include these values under the columns titled O.R. i.e. odds ratios.

Looking at the value of the odds ratio for price, our results show that keeping everything else constant, a unit increase in the average weekly basket price leads to a 5% reduction, on average, in the odds of choosing a particular grocery retailer. On the other hand, for every unit increase in the annual loyalty discount, the odds of choosing

[65] We follow the guidance of UCLA Statistical Consulting Group for interpreting coefficient values and corresponding odds ratios in logistic regression models which has been obtained from http://www.ats.ucla.edu/stat/mult_pkg/faq/general/odds_ratio.htm

a given retailer go up by 1%. The effect of increasing drive time to the grocery store is quite large in comparison to these two variables. For every additional minute of drive time the odds of store visit go down by 6%. The estimated coefficient and odds ratio values for the qualitative dummy coded variables can be interpreted as a marginal movement from the base level 'low' to either 'medium', 'high' or to the 'very high' level. The likelihood of the consumer selecting a particular retailer goes up significantly as the available product range in the store increases across these levels. A similar relationship is observed for the other qualitative variables. The estimates reveal a strong preference for very high levels of service whereby the odds of choosing a retailer go up 3 times when we move from low level of service to very high.

In addition to testing the statistical significance of individual parameters, we use the χ^2 likelihood ratio test as a measure for the goodness of fit of each respective model. This is computed automatically by *Stata*. The likelihood ratio test ("LR") is an important tool to assess the overall goodness of fit of a model that is based on using maximum likelihood estimation procedures (Louviere et al. 2000, pp. 53-55). This test is also useful when fitting the data to several model specifications because it allows the analyst to measure and compare the goodness of fit different model forms. We use the log-likelihood values obtained from the Stata output to manually compare structurally similar models to each other.

For an individual model, the test statistic is automatically calculated by *Stata* using the formula 2*('log-likelihood of constrained model' – 'log-likelihood at convergence') or alternatively as 2*('log-likelihood at convergence of less restrictive model' – 'log-likelihood at convergence more restrictive model'), with degree of freedom equal to the number of explanatory variables included in the model. The calculated test statistic is then compared to a critical value of the chi-square distribution with respect to the degrees of freedom. *Stata* automatically performs this statistical test when producing the output containing the estimated results. In the context of model CL1 we have very low *p*-value and reject the null hypothesis that the predictors included in the model have no explanatory power ($\chi^2(12)=245.24$).

Let us now consider model CL2 results, where we have included the 39 interaction terms we wanted to test for statistical significance. Out of these 39 interactions 7 were found to be significant.[66] These terms were: "Female x very high quality", "living in a household with 4 or more individuals x loyalty discount", "aged 18-44 x price", "purchases groceries at least 1/ month x price", "purchases groceries at least 1/ month x travel time to store", "does not own loyalty cards x loyalty discount", "does not own loyalty cards x travel time to store". We re-ran a new specification of the conditional logit and included these significant interaction terms. Following this process we eliminated a further three interaction terms. The model containing only significant interaction terms is titled model CL3 and includes the interaction terms: "Female x very high quality", "purchases groceries at least 1/ month x price", "purchases groceries at least 1/ month x travel time to store" and "does not own loyalty cards x loyalty discount". In interpreting the results, it should be noted that each of the interaction terms' corresponding coefficient values enter the model in an additive way. For example, the mean effect of a unit increase in the average basket price for individuals who shop online at least once a month, will be equal to the mean coefficient for price plus the mean coefficient of the interaction term "purchases groceries at least 1/ month x price". The results for this specification are presented in Table 3.9 below.

[66] We note that by clustering standard errors at the individual level the estimated standard errors are higher and thus there are fewer of significant interaction terms than without clustering. Nonetheless, by clustering standard errors in this way have obtained more reliable results.

Table 3.9 – Conditional logit model CL3

Variable	Coef.	|*z-stat*|	O.R.
Price	-0.068*** (0.008)	-8.82	0.934
Discount	0.011*** (0.002)	6.78	1.011
Time	-0.085*** (0.011)	-7.84	0.919
Medium Quality	0.551*** (0.124)	4.44	1.734
High Quality	0.554*** (0.142)	3.91	1.740
Very High Quality	1.152*** (0.209)	5.51	3.164
Medium Range	0.699*** (0.106)	6.63	2.013
High Range	0.917*** (0.128)	7.16	2.502
Very High Range	1.032*** (0.111)	9.29	2.807
Medium Service	0.862*** (0.110)	7.82	2.368
High Service	0.983*** (0.125)	7.83	2.673
Very High Service	1.169*** (0.147)	7.96	3.218
Female*VH Quality	-0.563*** (0.202)	-2.79	0.570
Frequent Online Shop*Price	0.038*** (0.011)	3.43	1.039
Frequent Online Shop*Time	0.069*** (0.016)	4.42	1.072
No Loyalty Cards*Discount	-0.011*** (0.003)	-3.25	0.989
Log-likelihood	-5083.889		
Nr. of Resp.	427		
Nr. of Obs.	18832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

Let us first compare the log-likelihoods at convergence for CL1 and CL3. There is an improvement in the overall significance of the model following the inclusion of covariates. The LR test outlined above can also be used to directly compare the goodness of fit between two models by testing the significance of the omitted predictors, 4 covariates in this case. The test statistic can be calculated as 2*(-5083.8893+5204.1841) = 120.295 at 4 degrees of freedom ($\chi^2(4)$= 120.295). This figure implies a rejection of the null hypothesis at the 1% level that the addition of the covariates does not improve the goodness of fit of the model.

In interpreting the interaction between the price variable and online shopping frequency, we note that preferences for basket price can to some degree be interpreted as a proxy for relative price sensitivity between groups of consumers.[67] In other words, a more negative coefficient for a given average basket price interaction term indicates greater price sensitivity relative to other groups. In this context, our results show that individuals who are frequent online shoppers of groceries products (i.e. shop online for groceries at least once a month) are relatively less price sensitive than those who never shop online or shop online only a few times a year. In addition those who prefer to shop online also care less about the travel time to the store. Furthermore, the results show that there are gender differences in preferences for very high quality products in store. We find that female shoppers care less about this attribute than male grocery shoppers when choosing between retailers. The results also show that households that tend *not* to participate in any loyalty schemes, care less about receiving a loyalty discount when choosing between grocery retailers. This is compared to those individuals who *do* participate in loyalty schemes and would prefer to have a loyalty discount offered by the grocery retailer. All the main explanatory variables and interaction terms in the results table are significant at the 1% level.

Having discussed the main results of the conditional logit model, we continue this section by presenting and discussing the main results obtained by running the mixed logit specifications. One of the considerations when fitting the data to the mixed logit model is to decide which coefficients should be fixed and which will be allowed to vary between individuals (Greene & Hensher 2003, Hole 2007a). The coefficients assumed to be random will capture preference heterogeneity in the data. This effect is contained in the estimated standard deviations of their respective mean coefficients. Statistically significant standard deviations suggest that the preferences exhibit heterogeneity for that attribute (i.e. variable). The analyst must also decide which *type* of distribution to impose on the respective attribute coefficients which are allowed to vary between respondents. The most commonly known, and widely used approach, applies the normal distribution to obtain the parameter values through maximum simulated log-likelihood estimation ("MSLE").

A key feature of the mixed logit model is that it relies on simulation methods. This is discussed at greater length in the methodological chapter when the choice of

[67] This interpretation of the price variable should be considered with caution as this is merely a proxy.

econometric model is discussed. One of the assumptions that the researcher must consider when running the mixed logit specification, is the number of draws to assume. The main trade-off between using low versus a high number of draws is that as the number of draws goes up, the accuracy of estimates increases at the expense of increased computational time (Hole 2007). Following Hole (2007), when trialling different model specifications we began the process by using only 50 Halton draws, which is the lowest recommended number. For comparative purposes, some of these results are presented in the Appendix, for example model ML2 (c). It is worth noting that the literature suggests that at 500 draws, coefficients converge to their true value (Hole 2007, Greene & Hensher 2003). The mixed logit specification results presented below have all be estimated using 500 Halton draws.

Let us first consider mixed logit models ML1 and ML2 where we have applied the normal distribution to the coefficients assumed to be random when running the specifications in *Stata*. These results are presented in Table 3.10 below. To test all attributes for presence of preference heterogeneity, model ML1 assumes that all the variable coefficients are distributed randomly following a normal distribution. In this specification, not all of the estimated standard deviations were found to be significant and alternative specifications were run in *Stata*. Model ML2 presents the results of the model with only significant standard deviations. Preferences were found to vary among households for the average basket price, loyalty discount, travel time, very high quality, very high range and high levels of service. The *p*-values for these standard deviations are small suggesting we can reject the null hypothesis of all standard deviations being equal to zero (Hole 2007).

Table 3.10 – Mixed logit models ML1 & ML2

	ML1			ML2		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev.**
Price	-0.091*** (0.009)	10.5	-0.068*** (0.008)	-0.090*** (0.009)	9.62	0.067*** (0.009)
Discount	0.013*** (0.002)	5.34	-0.027*** (0.003)	0.011*** (0.002)	4.91	0.024*** (0.003)
Time	-0.110*** (0.015)	7.47	0.118*** (0.015)	-0.102*** (0.014)	7.25	0.106*** (0.015)
Medium Quality	0.516*** (0.163)	3.17	-0.500*** (0.180)	0.475*** (0.158)	3.00	-
High Quality	0.603*** (0.199)	3.03	0.423*** (0.202)	0.573*** (0.193)	2.98	-
Very High Quality	0.864*** (0.195)	4.44	0.750*** (0.192)	0.840*** (0.183)	4.58	0.696*** (0.187)
Medium Range	0.758*** (0.143)	5.31	-0.020 (0.113)	0.740*** (0.135)	5.49	-
High Range	1.008*** (0.170)	5.92	-0.002 (0.116)	0.984*** (0.164)	6.02	-
Very High Range	1.206*** (0.161)	7.49	0.947*** (0.139)	1.201*** (0.157)	7.65	0.882*** (0.142)
Medium Service	1.156*** (0.161)	7.2	-0.342 (0.226)	1.127*** (0.154)	7.31	-
High Service	1.364*** (0.166)	8.22	0.752*** (0.180)	1.340*** (0.161)	8.33	0.721*** (0.162)
Very High Service	1.484*** (0.212)	7.00	0.369 (0.225)	1.403*** (0.203)	6.91	-
Log-likelihood	-4627.524			-4647.902		
Nr. of Resp.	427			427		
Nr. of Obs.	18,832			18,832		

Notes: Robust standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

In the analysis of results we refer to ML2 estimates where estimated coefficient and standard deviation values were found to be significant at the 1% level. Following Train (2009, pp. 149) we interpret the mean coefficient estimates using the cumulative standard normal distribution given by $\Phi(-b_k/s_k)$ with mean b_k and standard deviation s_k. By applying this approach, we find that individuals have preferences for lower prices. Using the above formula and with reference to a *z*-table, we determine that 9% of the distribution in preferences for the average basket price is above zero and 91% is below zero. This means that lower average basket prices are an attractive grocery retailer feature for 92% of consumers while 8% of consumers choose to shop at more expensive retailers and/ or are not insensitive to the higher price. On the other hand, 68% of consumers prefer choosing a grocery retailer that offers a loyalty rewarding scheme and corresponding discount, while 32% of consumers do not care about the discount aspect of the grocery retailer's offering. In terms of travel time to the store, 83% of consumers have a preference for a grocery retailer at a closer proximity, while just 17% are willing and/ or indifferent about travel further distances to a grocery store. Almost all consumers would prefer to frequent stores that offer a high level of service and very high quality products. 97% of households are attracted to stores with high service levels. In terms of service levels, 91% of households prefer stores with very high levels of product range and 89% households are attracted to stores with products that are of very high quality.

We re-ran the above mixed logit model by including the 39 interaction terms that were tested as part of the conditional logit model specification. Of these 39 interactions, 8 were found to be statistically significant. We then ran another specification including only these significant interaction terms. Results for both of these models, which have been labelled ML3 and ML4 respectively, are presented in Table 3.11 below.

Table 3.11 – Mixed logit models ML3 & ML4

	ML3			ML4		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev.**
Price	-0.059*** (0.023)	-2.61	0.051*** (0.009)	-0.081*** (0.014)	6.01	0.058*** (0.008)
Discount	0.023*** (0.007)	3.32	0.024*** (0.003)	0.012*** (0.003)	4.83	0.025*** (0.003)
Time	-0.150*** (0.035)	-4.25	0.096*** (0.017)	-0.129*** (0.016)	8.3	0.103*** (0.014)
Medium Quality	0.502*** (0.158)	3.18	-	0.489*** (0.159)	3.07	
High Quality	0.602*** (0.196)	3.08	-	0.588*** (0.196)	3.01	
Very High Quality	1.101*** (0.245)	4.5	0.652*** (0.174)	1.161*** (0.243)	4.78	0.675*** (0.187)
Medium Range	0.738*** (0.132)	5.57	-	0.737*** (0.134)	5.51	
High Range	0.978*** (0.161)	6.08	-	0.982*** (0.163)	6.01	
Very High Range	1.110*** (0.274)	4.06	0.923*** (0.131)	1.029*** (0.171)	6.00	0.897*** (0.134)
Medium Service	1.137*** (0.156)	7.31	-	1.136*** (0.156)	7.29	
High Service	1.110*** (0.177)	6.27	0.719*** (0.143)	1.260*** (0.169)	7.47	0.701*** (0.147)
Very High Service	1.426*** (0.206)	6.91	-	1.418*** (0.206)	6.87	-
Female*Price	-0.017 (0.017)	-1.05		-	-	-
Female*Discount	0.002 (0.005)	0.55		-	-	-
Female*Time	-0.033 (0.028)	-1.16		-	-	-
Female*VH Quality	-0.579** (0.238)	-2.44		-0.571** (0.237)	-2.41	
Female*VH Range	0.027 (0.226)	0.12		-	-	-
Female*High Service	0.202 (0.200)	1.01		-	-	-
Large Household*Price	0.024 (0.015)	1.61		-	-	-
Large Household*Discount	0.008* (0.005)	1.72		0.009* (0.005)	1.86	
Large Household*VH Range	-0.182 (0.224)	-0.81		-	-	-
Unemployed*Price	-0.019 (0.018)	-1.08		-	-	-
Student*Price	0.016 (0.014)	1.14		-	-	-
Unemployed*Discount	-0.002 (0.009)	-0.18		-	-	-
Student*Discount	-0.002 (0.005)	-0.31		-	-	-
Unemployed*Time	0.021 (0.042)	0.5		-	-	-
Student*Time	-0.038	-1.1		-	-	-

	(0.035)				
No Car*Time	-0.003 (0.026)	-0.11	-	-	-
No Car*VH Range	0.463** (0.230)	2.01	0.421* (0.250)	1.68	-
18 – 44 Age Group*Price	-0.027 (0.019)	-1.45	-	-	-
18 – 44 Age Group*Discount	-0.002 (0.005)	-0.39	-	-	-
< £22,000 HI*Price	-0.025* (0.015)	-1.65	-0.038*** (0.014)	-2.79	-
> £45,000 HI*Price	0.006 (0.031)	0.2	-	-	-
> £45,000 HI*Discount	-0.006 (0.006)	-0.87	-	-	-
> £45,000 HI*VH Quality	0.377 (0.293)	1.29	-	-	-
> £45,000 HI*High Service	0.530** (0.235)	2.26	0.374* (0.228)	1.64	-
< £22,000 HI*Discount	-0.008 (0.005)	-1.61			-
> £45,000 HI*Time	0.043 (0.043)	1.01	-	-	-
Frequent Online Shop*Price	0.050*** (0.013)	3.82	0.044*** (0.014)	3.14	
Frequent Online Shop*Discount	0.000 (0.005)	-0.05	-	-	-
Frequent Online Shop*Time	0.090*** (0.023)	3.89	0.088*** (0.023)	3.89	
Infrequent Online Shop*Price	-0.012 (0.014)	-0.83	-	-	-
Infrequent Online Shop*Discount	-0.004 (0.005)	-0.81	-	-	-
Infrequent Online Shop*Time	0.034 (0.024)	1.41	-	-	-
Infrequent Online Shop*VH Range	-0.061 (0.239)	-0.25	-	-	-
No Loyalty Cards*Price	0.009 (0.016)	0.57	-	-	-
No Loyalty Cards*Discount	-0.017*** (0.006)	-2.86	-0.017*** (0.004)	-3.96	
No Loyalty Cards*Time	0.051 (0.032)	1.59	-	-	-
1-2 Loyalty Cards*Time	0.016 (0.026)	0.59	-	-	-
1-2 Loyalty Cards*Discount	-0.006 (0.005)	-1.34	-	-	-
1-2 Loyalty Cards*Price	-0.010 (0.018)	-0.58	-	-	-
Log-likelihood	-4527.994		-4574.406		
Nr. of Resp.	427		427		
Nr. of Obs.	18832		18832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

The four interaction terms found to be significant in the conditional logit model specification, are also significant when fitting the data to the mixed logit model. In addition, the mixed logit model represents an improvement over the conditional logit specification as it picked up additional, statistically significant variation in the data. In addition to producing larger estimated coefficient values for the attributes, the above results contain a further 4 significant interactions compared to CL3. The statistical significance of these terms varies between 1%, 5% and 10% level of significance. The additional covariates found to be significant by running the mixed logit model are: "living in a household of 4 or more individuals x discount", "not driving a car to go grocery shopping x very high range", "household income under £22,000 x price" and "household income above £45,000 x high level of service".

The two latter interactions capture differences in preferences between higher and lower income groups. Shoppers who have lower disposable household incomes, prefer lower prices and are likely to be more sensitive to price when choosing between retailers. On the other hand individuals who enjoy higher disposable household incomes care less about price and prefer having a high level of service in store when choosing between retailers. In addition, large households prefer a grocery retailer who offers a loyalty discount and those who do not drive a car to go grocery shopping prefer choosing a grocery retailer who offers a very high range of products in store.

The distribution of preferences implied by the estimated standard deviations are similar to what we observe in model ML2. 92% of respondents prefer a grocery retailer who charges lower average basket prices while 69% of respondents prefer to shop at a store where they can obtain a loyalty discount. In terms of store proximity, 89% of respondents prefer to travel shorter distances to the grocery store. While 96% of individuals prefer grocery retailers who offer very high quality of products, 87% prefer very high range of products in store and 96% prefer high levels of service at the grocery store. Considering the above, the loyalty scheme attribute exhibits the most variation in preferences between individuals in the sample.

We refer to the LR test to compare the goodness-of-fit between models ML2 (without interactions) and ML4 (with significant interactions). At 8 degrees of freedom we have $\chi^2(8)= 73.496$. This figure is greater than the critical value at 8 d.f. and therefore implies a rejection of the null hypothesis at the 1% level that the addition of the covariates does not improve the goodness of fit of the model. We now consider alternative distributional assumptions accommodated by the mixed logit model which may be imposed on the variables assumed to be random instead of the normal distribution.

The mixed logit models discussed above assumed the normal distribution for the coefficients. However, there are other distributions which can also be applied to the modelling. Other less well known distributions, also more rarely discussed in the literature, are the uniform, triangular and lognormal (Greene & Hensher 2003). The uniform and triangular distributions will generally report the same or similar values for parameter estimates as when assuming the normal (Greene & Hensher 2003). In contrast, the imposition of a lognormal distribution can lead to different estimates and create additional challenges as compared to assuming simply normally distributed coefficients. Namely, problems relating to non-convergence and complexities associated with the calculation of willingness-to-pay estimates (Greene & Hensher 2003). However, the main benefit of the lognormal distribution is that it restricts coefficient values to be either positive or negative for all individuals which may improve the results of the model in some cases. This may be desirable for variables such as price, which is likely to always have a negative coefficient value for individuals.

This type of distributional assumption may be desirable if the researcher knows for certain whether a change in the value of the attribute, will either increase or decrease total utility for that absolute vast majority of consumers within the population. For example, the researcher can say for certain that the costs incurred of choosing a given alternative will always have a negative impact on utility (Train 2009, pp. 149-150). The evidence overall is mixed with regards to the benefits of the lognormal because in some cases it will improve *or* reduce the goodness of fit of the data (Hole 2008). In this context, the researcher must decide which assumptions are appropriate for their particular data by comparing the results of different model specifications. The *mixlogit* command in *Stata 12* allows the analyst to either impose a normal or lognormal distribution by adjusting the underlying command syntax and adjusting the variable assumed to be log-normally distributed (Hole 2007).

When fitting the data to the mixed logit model we wanted to compare if restricting the sign of some of the coefficient values, either being positive or negative, would improve the quality and explanatory power of our results. In doing so, we impose an assumption in the modelling that both the average weekly basket price and travel time to the store represent a direct cost to the consumer and subtract from gained utility. With respect to the other variables in the data, it was not obvious whether they would result either only positive or negative coefficient values for all respondents and were therefore assumed to follow a normal distribution as in the other mixed logit specifications. Table 3.12 below presents the results of two mixed logit specifications which assume that price and time are randomly distributed following a log-normal distribution. Model ML5 contains only the main explanatory variables while model ML6 contains 11 significant interaction terms. The results for the model which restricts time and price to be log-normally distributed and contains the 39 candidate interactions terms is labelled model ML6 (a) and can be found in the Appendix under Table A.3.7.

Table 3.12 – Mixed logit models ML5 & ML6

	ML5			ML6		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev.**
Price	-0.120*** (0.020)	5.87	0.180*** (0.063)	-0.127*** (0.021)	10.79	0.135*** (0.062)
Discount	0.014*** (0.003)	5.55	0.024*** (0.003)	0.017*** (0.003)	6.04	0.024*** (0.003)
Time	-0.178*** (0.041)	4.32	-	-0.124*** (0.026)	8.78	0.194*** (0.064)
Medium Quality	0.441*** (0.160)	2.76	-	0.439*** (0.166)	2.64	-
High Quality	0.537*** (0.191)	2.81	-	0.534*** (0.199)	2.68	-
Very High Quality	0.878*** (0.187)	4.7	0.607** (0.244)	1.194*** (0.256)	4.67	0.598*** (0.193)
Medium Range	0.710*** (0.134)	5.3	-	0.685*** (0.134)	5.1	-
High Range	0.945*** (0.160)	5.91	-	0.921*** (0.164)	5.63	-
Very High Range	1.145*** (0.156)	7.36	0.901*** (0.131)	0.947*** (0.175)	5.4	0.874*** (0.135)
Medium Service	1.090*** (0.154)	7.09	-	1.124*** (0.155)	7.26	-
High Service	1.298*** (0.156)	8.31	0.720*** (0.153)	1.194*** (0.173)	6.89	0.683 (0.145)
Very High Service	1.416*** (0.204)	6.93	-	1.405*** (0.208)	6.77	-
Female*Time				-0.047** (0.024)	1.96	-
Female*VH Quality				-0.595** (0.244)	2.44	-
Large Household*Price				0.025** (0.012)	2.02	-
Unemployed*Price				-0.028** (0.014)	2.04	-
No Car*VH Range				0.443* (0.247)	1.79	-
< £22,000 HI*Price				-0.025* (0.014)	1.83	-
> £45,000 HI*Price				0.040** (0.019)	2.09	-
> £45,000 HI*High Service				0.464** (0.233)	1.99	-
Frequent Online Shop*Price				0.039*** (0.014)	2.9	-
Frequent Online Shop*Time				0.090*** (0.019)	4.77	-
No Loyalty Cards*Discount				-0.018*** (0.005)	3.65	-
Log-likelihood	-4656.764			-4538.897		
Nr. of Resp.	427			427		
Nr. of Obs.	18832			18832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

The results table shows that all variables in ML5 are significant at the 1% level. Compared to model ML2, where all coefficients followed a normal distribution, coefficients for time and price are restricted to being only negative in the population for model ML5. This assumption has had an impact on the magnitudes of the mean coefficients for price and time which are now visibly more negative in model ML5 than ML2. In addition, the variable representing the loyalty scheme discount has a relatively larger coefficient value than in specification ML2. However, if we compare the log-likelihood values at convergence between ML2 and ML6, the normal distribution assumption for all the coefficients results in a slightly better fitting model. This result is only indicative, however, the fact that the normal distribution results in a slightly better fitting model is consistent with the evidence from the literature (Greene & Hensher 2003, Hole 2008).

Considering model ML6 results, the addition of interactions results in more negative coefficient values for price and time and a bigger positive coefficient value for the discount variable. In addition, the interaction terms found to be significant are not strictly that same as in the previous specification ML4 and these also vary in their statistical significance levels. When running model ML6, the results pick up gender differences in preferences for travel time to the store. In other words, we find that the effect of being female leads to a more negative coefficient for the time attribute indicating that women shoppers are more sensitive to the location of the store than their male counterparts. In these results, the covariate "household income is greater than £45,000 x price" is statistically significant at the 5% level, further evidencing the differences in price sensitivity between households earning meaningfully different disposable incomes. The interaction between the lowest income category and price is also significant but only at the 10% level.

We also find that respondents who indicated they were unemployed prefer grocery retailers who offer lower average basket prices indicating these individuals are likely to be more price sensitive than individuals in other occupations. In this model, the interaction between large household size and discount is no longer significant. Instead, the interaction between large household size (i.e. 4 or more people) and price is significant indicating that large households have a preference for lower average basket prices when choosing between grocery retailers and are likely to be more price sensitive. Applying the LR test to compare models ML5 and ML6, we can reject the null hypothesis that the addition of 11 interaction terms does not constitute an improvement over the more restricted model ($\chi^2(11)= 117.867$).

As noted at the beginning of this section, we place the most weight on models ML2 and ML4 with 8 interaction terms in the context of the discussion of results. These models represent an improvement over the conditional logit specifications by accommodating unobservable (i.e. not tied to covariates) preference heterogeneity between individuals captured in the standard deviations of coefficients. We note that these models assume the normal distribution, which according to the literature is the *least* restrictive assumption to impose on the coefficients assumed to be random. Further, the preferred specifications apply 500 Halton draws, include population weights and cluster standard errors at the individual level. Before proceeding to the discussion section, we present additional results obtained using the post-estimation commands for the *mixlogit* in *Stata 12*, including willingness-to-pay and individual-level parameter estimates.

3.5 Willingness-to-Pay & Individual-Level Parameter Estimation

In this section we present willingness-to-pay estimates for different grocery retailer attributes. Further below we include some graphical representations of the distributions of preferences using individual-level mean coefficients estimates. The purpose of computing willingness-to-pay estimates is to provide a sense of ordered preferences for the attributes included in this DCE. The WTP estimates were calculated using estimates obtained by using the estimates from the conditional logit models CL1 and CL3 and by running additional mixed logit model specifications with fixed price coefficients. The analysis of WTP estimates, offers an additional means to interpret the results by identifying the relative importance of attributes in an ordered and simple way (Hole & Kolstad 2012).

In the context of the below results, we highlight that our experiment was not designed to produce precise WTP estimates. This task would have required very specific considerations in the survey design in terms of carefully designed variation in price levels displayed to respondents between choice situations. As such, the estimates of the WTP for various attributes should be interpreted with caution. The results in this section have been derived by applying a commonly used approach that requires running a model specification using a fixed cost coefficient (Hole & Kolstad 2012). While it may be unreasonable to assume that the effect of price is the same for all individuals, using a fixed monetary cost coefficient is convenient as it overcomes many modelling issues in random parameter models such as ours (Bliemer & Rose 2013, Hole & Kolstad 2012). In current literature there are other, newer and more advanced approaches for modelling WTP. However, these approaches are beyond the focus and scope of this paper.

Following the approach outlined above, we assume that attribute $\boldsymbol{k}$ has a fixed coefficient equal to $\boldsymbol{\beta_k}$. In the case of our data, we assume a fixed cost coefficient $\boldsymbol{\beta_c}$, which is the average basket price. We can then express willingness-to-pay as: $\boldsymbol{WTP_k = -\beta_k/\beta_c}$. The below Table 3.13 presents the WTP estimates computed on the basis of model CL1 and CL3 coefficient estimates.

Table 3.13 – Models CL1 & CL3 Willingness-to-Pay Estimates (95% Confidence Intervals)

	CL1			**CL3**		
Grocery Retailer Attribute	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**
Grocery Basket Price	-	-	-	-	-	-
Loyalty Scheme Discount	0.17	0.11	0.23	0.16	0.10	0.22
Travel Time to Store (mins)	-1.13	-1.43	-0.82	-1.24	-1.55	-0.93
Medium Quality	10.26	5.93	14.58	8.06	4.41	11.71
High Quality	10.20	5.15	15.25	8.11	3.88	12.34
Very High Quality	15.26	9.53	20.99	16.86	9.88	23.84
Medium Range	13.00	8.87	17.13	10.24	6.84	13.63
High Range	17.00	11.78	22.22	13.42	9.08	17.76
Very High Range	19.27	14.29	24.25	15.11	10.97	19.24
Medium Service	15.54	10.72	20.36	12.62	8.54	16.69
High Service	18.18	11.93	24.44	14.39	9.16	19.62
Very High Service	21.19	14.87	27.51	17.11	11.58	22.63

We must note that many of the above WTP estimates, particularly for the qualitative attributes in models Cl1 and CL2, appear quite inflated, however, we interpret their magnitudes relative to each other. This allows us to get a sense of ordered preferences individuals may have for these attributes when choosing between retailers. Looking at the estimates for CL1, in terms of ordered preferences, when choosing between retailers, receiving very high levels of service is the most valuable attribute to shoppers. Based on the above results, consumers are willing to pay £0.17 more for their weekly average basket price in return for an additional unit increase in the average annual loyalty discount they receive. The negative willingness-to-pay estimates for the store proximity variable are the result of this variable's negative coefficient estimates which indicate that increasing the value of this variable (i.e. increasing drive time) decreases consumer utility. In this context, consumers are willing to pay £1.13 more for their average weekly basket price if they must travel one minute less to the store by car.

Let us now consider the mixed logit specifications we ran to compute additional WTP estimates. In the case of the mixed logit model, the distribution for WTP is equal to the variable's assumed distribution, scaled with respect to the monetary fixed cost coefficient. Using the *wtp* post-estimation command on *Stata 12*, we specified the default 'delta method' to determine the confidence intervals for the mean WTP values.[68] This approach assumes that WTP is normally distributed and *Stata* automatically calculates the relevant confidence intervals for the estimated values. This implies that WTP will also likely be normally distributed. Furthermore, the assumption that WTP is normally distributed is said to hold when the sample size is large the cost coefficient is a precise estimate (Hole 2007b).

We re-ran a number of different mixed logit specifications and in each of these, assumed a fixed coefficient value for the average basket price coefficient. In total, we ran a further 4 specifications for the purposes of WTP estimates: ML7 (a), ML7 (c), ML8 (a) and ML8 (b). The results for these models are presented under Tables 3.14 and 3.16 below.

[68] Please refer to Hole (2007b) and Bliemer & Rose (2013) for an appraisal of the different methods available for calculating confidence intervals for willingness-to-pay measures

Table 3.14 – Mixed logit models ML7 (a) & ML8 (a)

	ML7 (a)			ML8 (a)		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev.**
Price	-0.061***	-9.40		-0.060***	-9.32	
	(0.006)		-	(0.006)		-
Discount	0.010***	4.41	0.025***	0.011***	4.64	0.025***
	(0.002)		(0.003)	(0.002)		(0.003)
Time	-0.095***	-7.52	0.109***	-0.133***	-5.9	0.301**
	(0.013)		(0.013)	(0.023)		(0.124)
Medium Quality	0.534***	3.79		0.533***	3.79	-
	(0.141)		-	(0.140)		
High Quality	0.615***	3.52		0.605***	3.48	-
	(0.175)		-	(0.174)		
Very High Quality	0.838***	4.74	-0.684***	0.826***	4.69	0.646***
	(0.177)		(0.210)	(0.176)		(0.210)
Medium Range	0.772***	5.83		0.732***	5.58	-
	(0.132)		-	(0.131)		
High Range	1.007***	6.47		0.962***	6.27	-
	(0.156)		-	(0.154)		
Very High Range	1.295***	8.60	1.016***	1.275***	8.57	1.004**
	(0.151)		(0.116)	(0.149)		(0.110)
Medium Service	1.061***	7.20		1.023***	6.97	-
	(0.147)		-	(0.147)		
High Service	1.382***	8.70	0.726***	1.342***	8.55	0.739***
	(0.159)		(0.131)	(0.157)		(0.127)
Very High Service	1.420***	7.17		1.396***	7.13	-
	(0.198)		-	(0.196)		
Log-likelihood	-4780.929			-4780.832		
Nr. of Resp.	427			427		
Nr. of Obs.	18,832			18,832		

Notes: Robust standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

Model ML7 (a) is a mixed logit model without interactions, which assumes that the attributes assumed to be random follow a normal distribution. On the other hand, Model ML8 (a) is also a mixed logit, however the time attribute is assumed to follow a random log-normal distribution. The estimated values are similar between these two models, bar the coefficient for time travelling to the store. Below, in Table 3.15 we present the WTP estimates derived from the above results. The WTP values we obtain below are very similar to those based on estimates from the conditional logit model specifications. However, the attribute representing travel time to the store represents a greater value. Based on the WTP estimates below individuals are willing to pay an additional £2.23 for their basket price to travel one minute less to the store. On the basis of the relative WTP values presented below, both the level of product range and levels of service are very important to shoppers choosing between grocery stores.

Table 3.15 – Models ML7 (a) & ML8 (a) Willingness-to-Pay Estimates (95% Confidence Intervals)

	ML7 (a)			ML8 (a)		
eGrocery Retailer Attribute	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**
Grocery Basket Price	-	-	-	-	-	-
Loyalty Scheme Discount	0.17	0.09	0.25	0.18	0.09	0.26
Travel Time to Store (mins)	-1.57	-1.98	-1.15	-2.23	-2.45	-1.49
Medium Quality	8.77	4.32	13.21	8.92	4.38	13.46
High Quality	10.11	4.53	15.69	10.13	4.44	15.83
Very High Quality	13.77	8.13	19.42	13.84	8.09	19.59
Medium Range	12.68	8.30	17.07	12.25	7.81	16.69
High Range	16.54	11.26	21.81	16.12	10.77	21.46
Very High Range	21.26	16.00	26.53	21.35	15.98	26.71
Medium Service	17.43	11.91	22.96	17.13	11.60	22.66
High Service	22.69	15.94	29.45	22.47	15.71	29.24
Very High Service	23.32	16.15	30.48	23.39	16.14	30.63

Table 3.16 below presents the results for models ML7 (c) and ML8 (b) which assume a fixed coefficient for the average basket price. ML7 (c) is an extension of model ML7 (a) as it contains interaction terms with the remaining assumptions remaining constant. Similarly, model ML8 (b) represents an extension of model ML (a), as it follows the same assumptions albeit it contains interaction terms. Table 3.17 further below presents the WTP estimates computed on the basis of the results of the above models.

Table 3.16 – Mixed logit models ML7 (c) & ML8 (b)

	ML7 (c)			ML8 (b)		
Variable	**Coef.**	**|z-stat|**	**St. Dev.**	**Coef.**	**|z-stat|**	**St. Dev.**
Price	-0.058*** (0.010)	-5.78	-	-0.057*** (0.010)	-5.71	-
Discount	0.013*** (0.003)	5.03	0.024*** (0.003)	0.013*** (0.003)	5.09	0.024*** (0.003)
Time	-0.122*** (0.015)	-8.35	0.102*** (0.014)	-0.126*** (0.033)	-3.78	0.316* (0.239)
Medium Quality	0.534*** (0.145)	3.67	-	0.533*** (0.145)	3.68	-
High Quality	0.623*** (0.181)	3.44	-	0.612*** (0.180)	3.39	-
Very High Quality	1.231*** (0.243)	5.06	0.616*** (0.184)	1.207*** (0.240)	5.02	0.567*** (0.198)
Medium Range	0.753*** (0.131)	5.75	-	0.715*** (0.129)	5.52	-
High Range	0.987*** (0.155)	6.35	-	0.948*** (0.154)	6.15	-
Very High Range	1.276*** (0.150)	8.51	0.992*** (0.121)	1.248*** (0.149)	8.35	0.984*** (0.112)
Medium Service	1.077*** (0.149)	7.23	-	1.049*** (0.148)	7.07	-
High Service	1.291*** (0.166)	7.8	0.709*** (0.129)	1.265*** (0.168)	7.54	0.693 (0.138)
Very High Service	1.421*** (0.199)	7.13	-	1.396*** (0.198)	7.04	-
Female*Time	-	-	-	-0.048* (0.024)	-2.02	-
Female*VH Quality	-0.672*** (0.243)	-2.76	-	-0.661*** (0.241)	-2.75	-
< £22,000 HI*Price	-0.035*** (0.011)	-3.28	-	-0.035*** (0.011)	-3.26	-
> £45,000 HI*High Service	0.358* (0.209)	1.71	-	0.370* (0.208)	1.78	-
Frequent Online Shop*Price	0.035*** (0.011)	3.13	-	0.035*** (0.011)	3.14	-
Frequent Online Shop*Time	0.085*** (0.022)	3.89	-	0.076*** (0.017)	4.43	-
No Loyalty Cards*Discount	-0.018*** (0.005)	-3.65	-	-0.018*** (0.005)	-3.61	-
Log-likelihood	-4672.343			-4661.221		
Nr. of Resp.	427			427		
Nr. of Obs.	18,832			18,832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

Table 3.17 – Models ML7 (c) & ML8 (c) Willingness-to-Pay Estimates (95% Confidence Intervals)

	ML7 (c)			ML8 (b)		
Grocery Retailer Attribute	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**	**Average WTP (£)**	**Lower bound (£)**	**Upper bound (£)**
Grocery Basket Price	-	-	-	-	-	-
Loyalty Scheme Discount	0.22	0.11	0.34	0.23	0.12	0.35
Travel Time to Store (mins)	-2.11	-2.85	-1.37	-2.22	-2.51	-1.62
Medium Quality	9.21	3.74	14.67	9.36	3.79	14.94
High Quality	10.74	4.19	17.29	10.75	4.06	17.45
Very High Quality	21.24	11.05	31.43	21.19	10.82	31.56
Medium Range	12.99	7.36	18.63	12.56	6.90	18.22
High Range	17.03	9.97	24.09	16.64	9.51	23.77
Very High Range	22.01	14.19	29.84	21.92	14.00	29.84
Medium Service	18.58	10.68	26.48	18.42	10.51	26.33
High Service	22.27	12.97	31.57	22.22	12.73	31.71
Very High Service	24.51	13.93	35.09	24.52	13.83	35.20

The above WTP estimates are similar to the results when running the other model specifications. Service levels at the grocery store and both product quality and range are valuable retailer attributes to grocery shoppers, with high levels of service being the most valuable on the basis of ordered preferences. We note that the time attributes in the above models follow two different distributions: normal and log-normal respectively. Nonetheless, these specifications have produced relatively similar WTP estimates for the time attribute. In the above table, the loyalty discount has a larger magnitude relative to the previous models ML7 (a) and ML8 (a) as a result of the inclusion of interaction terms. We continue this section by considering the individual-level parameter estimates we derived using results from mixed logit models ML2 and ML4.

This additional feature of the mixed logit model allows the researcher to graphically map the distributions of preferences for the attributes using individual-level parameter estimates via simulation. In performing this procedure, the researcher must obtain mean parameter estimates for each sampled observation. In other words, perform simulations of the data on the basis of the sequence of choices made by respondent *n* and compute the individual-level coefficient estimate for each respondent. These individual-level parameter estimates can then be mapped using kernel density plots (Greene. & Hensher 2003). This exercise can further assist in the interpretation of results through a graphical representation of the distribution of preferences as they appear in the data (Train 2009 pp. 259-281). Let us consider the details underlying this post-estimation technique.

The values for the individual-level coefficients $\boldsymbol{\beta}$ correspond to the sequences of choices made by individual respondents. The *mixlbeta* command enables the analyst to approximate the individual-level mean coefficients of the estimated variables with simulation procedures that use Halton draws. Thus instead of estimating a mean coefficient value $\boldsymbol{\beta_k}$ which applies across respondents as done previously, we are computing $\boldsymbol{\beta_n}$ which represents the individual-level parameter estimates. Following Hole (2013), the expected value of $\boldsymbol{\beta}$ is conditional on the pattern of choices $\mathbf{y_n}$ and the set of alternatives defined by their respective attributes in $\mathbf{x_n}$:

$$\boldsymbol{E}[\boldsymbol{\beta}|\boldsymbol{y_n},\boldsymbol{x_n}] = \frac{\int \boldsymbol{\beta} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{exp(x'_{njt}\beta)}{\sum_{j=1}^{J} exp(x'_{njt}\beta)}\right]^{y_{njt}} f(\boldsymbol{\beta}|\boldsymbol{\theta})}{\int \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{exp(x'_{njt}\beta)}{\sum_{j=1}^{J} exp(x'_{njt}\beta)}\right]^{y_{njt}} f(\boldsymbol{\beta}|\boldsymbol{\theta})}$$

Then $\boldsymbol{E}[\boldsymbol{\beta}|\boldsymbol{y_n},\boldsymbol{x_n}]$ can be approximated for individual $\boldsymbol{n}$ by taking $\boldsymbol{R}$ number of draws for each respondent using the distribution of $\boldsymbol{\beta}$ as follows:

$$\widehat{\boldsymbol{\beta_n}} = \frac{\frac{1}{R}\sum_{\boldsymbol{r=1}}^{\boldsymbol{R}} \boldsymbol{\beta_n^{[r]}} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp\left(x'_{njt}\beta_n^{[r]}\right)}{\sum_{j=1}^{J} \exp\left(x'_{njt}\beta_n^{[r]}\right)}\right]^{y_{njt}}}{\frac{1}{R}\sum_{r=1}^{R} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp\left(x'_{njt}\beta_n^{[r]}\right)}{\sum_{j=1}^{J} \exp\left(x'_{njt}\beta_n^{[r]}\right)}\right]^{y_{njt}}}$$

We obtained individual-level mean coefficient distributions by relying on the post-estimation procedures for *mixlogit* in *Stata 12* as set out in Hole (2007). Using the command *mixlbeta* we computed simulated individual-level coefficient estimates and applied the *kdenisty* command to generate the graphs. The approach we adopt follows the steps set out by Hole (2013) to estimate individual-level coefficients in *Stata 12* using the *mixlbeta* command. The process firstly requires estimation of a model specification. Then the values of individual-level coefficients and graphs are approximated by considering the sequence of choices made by an individual respondent and drawing from this distribution using the above commands. Only the variables assumed to be random can be used for this post-estimation procedure. The results are presented below in Figures 3.1 – 3.6.

Figure 3.1 – Individual-level parameter estimate for grocery basket price

ML2 – Grocery Basket Price ML4 – Grocery Basket Price

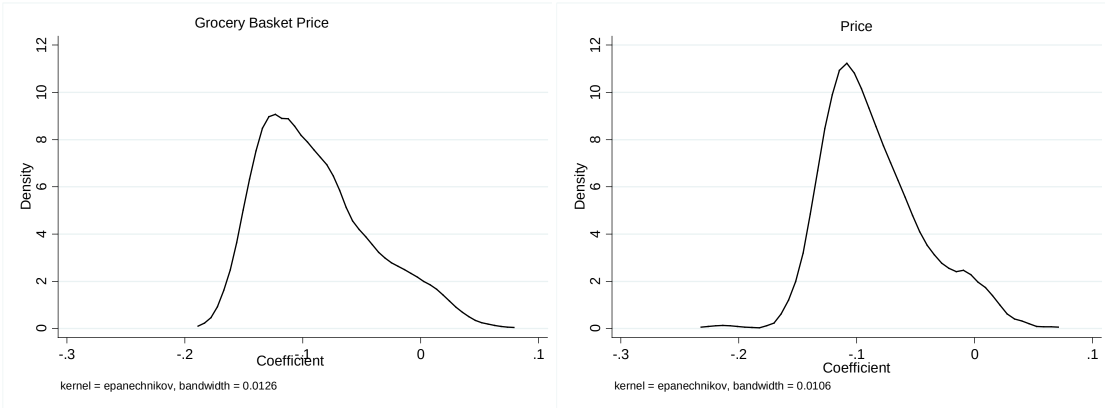


Figure 3.2 – Individual-level parameter estimate for loyalty scheme discount

ML2 – Loyalty Scheme Discount ML4 – Loyalty Scheme Discount

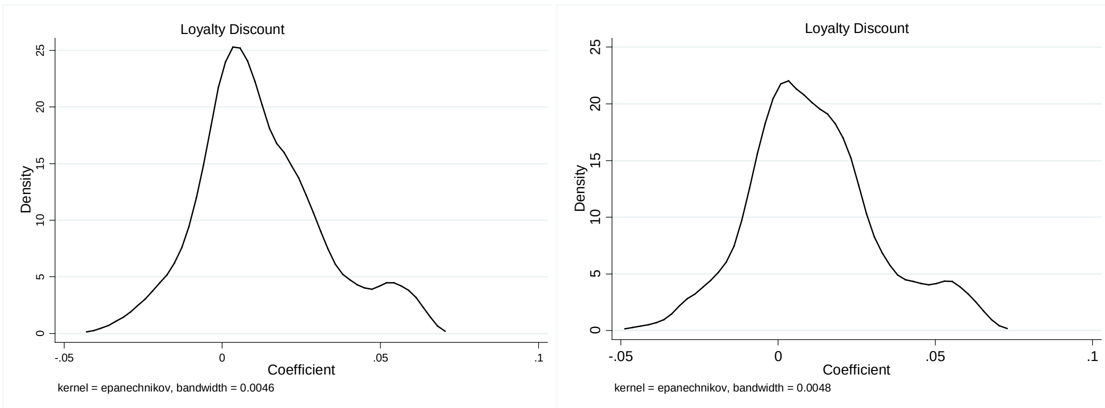
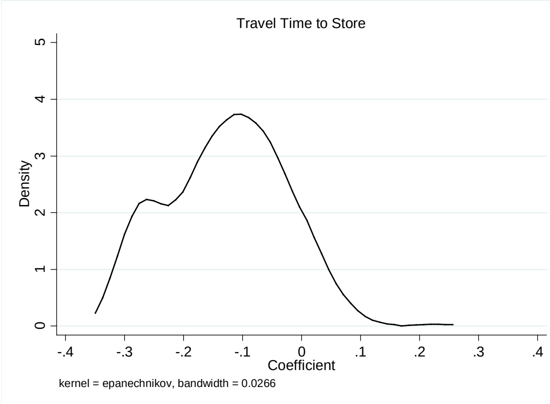


Figure 3.3 – Individual-level parameter estimate for travel time to store

ML2 – Travel Time to Store



ML4 – Travel Time to Store

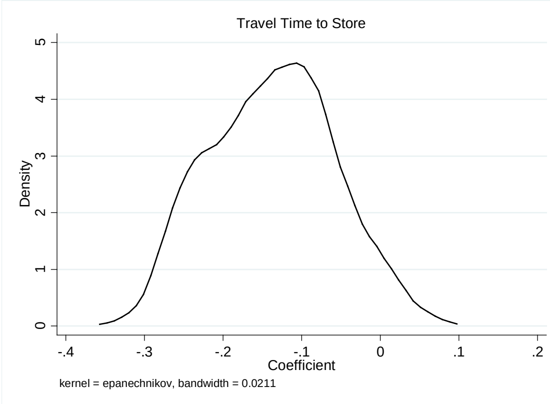
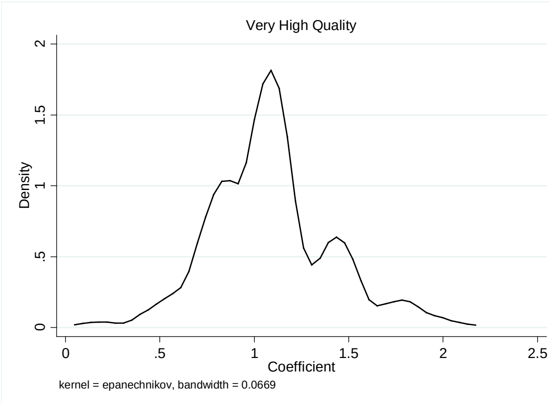


Figure 3.4 – Individual-level parameter estimate for very high quality

ML2 – Very High Quality



ML4 – Very High Quality

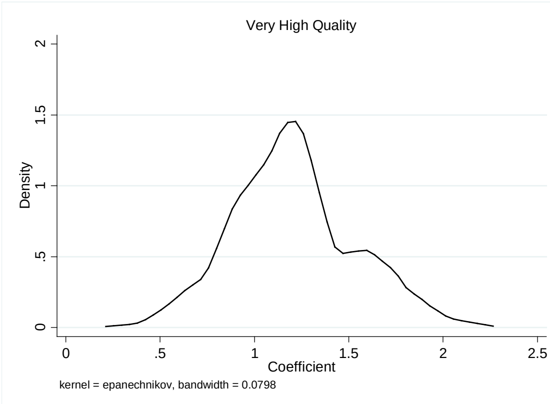
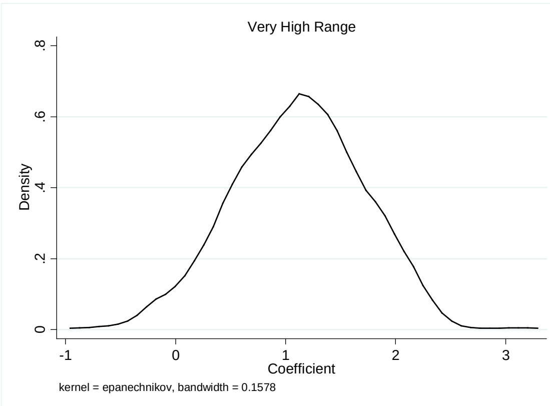


Figure 3.5 – Individual-level parameter estimate for very high range

ML2 – Very High Range



ML4 – Very High Range

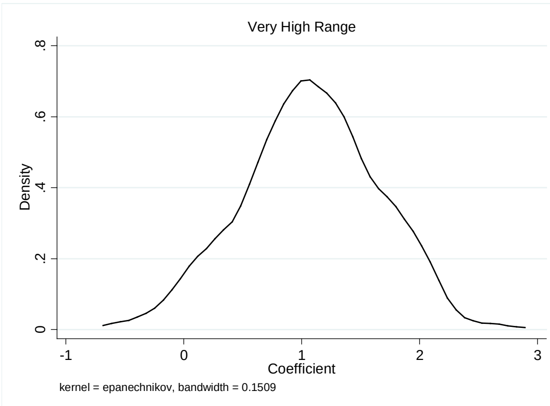
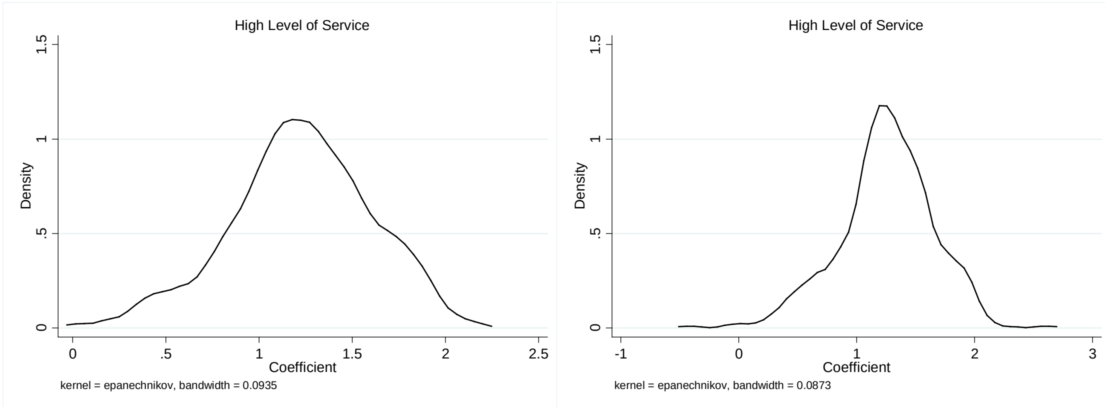


Figure 3.6 – Individual-level parameter estimate for high level of service

ML2 – High Level of Service   ML4 – High Level of Service



The above kernel density graphs representing the distributions of individual-level coefficient estimates are slightly different for the two model types. In Figure 3.6 above, model ML4 with interactions produces a distribution which is centred more closely around the mean value than in the case of ML2. However this is not consistent across the above figures. The distribution of preferences for travel time to the store appears to be more widely spread compared to the other distributions. Conversely, the distributions for very high quality are centred more closely around the mean, indicating the general preference for higher levels of product quality in store. However, these curves are uneven with some bumps due to unobservable variation in preferences. The loyalty scheme discount distribution follows a relatively smooth normal distribution in the context of model ML2 and is spread more widely after including the 8 interaction terms. The next section evaluates the results of the empirical work presented above in the context of the literature review chapter where we considered whether loyalty schemes could induce perceived switching costs in consumers choosing between retailers.

3.6 Discussion

In this section we evaluate our results in the context of evidence from real-world markets as well as the theoretical and empirical literature reviewed in the first chapter. We explain how our results may affect the assumptions on consumer preferences which enter theoretical models used to study markets for non-durable goods. We also outline potential competition policy implications of loyalty inducing strategies in retail markets on the basis of these results. In doing so we consider whether the trends we observe in our data substantiate the evidence on actual consumer behaviour in retail markets, in particular grocery retail. We briefly look at recent developments in the way that firms use loyalty rewarding schemes in practice and suggest directions for future research. We begin the discussion by presenting a summary of the empirical results of the discrete choice experiment with the most emphasis placed on parameter estimates obtained through models ML2 and ML4.

Individuals took part in our discrete choice experiment by completing a survey with 11 choice situations (survey questions) as well as sociodemographic questions. In each choice situation participants were presented with 4 alternative grocery retailer profiles and had to select only one as their preferred alternative. Each choice situation presented respondents with different trade-off between various combinations of retailer features, including different sizes of loyalty scheme discount, different average basket prices, alternative travel time to the store and 4 levels of product quality, product range and customer service respectively. The choice of what values to present to respondents to ensure visible trade-offs between alternatives is explained at length in the methodological chapter where we discuss survey design. The choice situations presented to respondents assumed that two of the four grocery retailers offered a loyalty scheme, while the others did not. Through the sequence of choices made by each individual respondent we were able to measure the relative importance of grocery retailer attributes to consumers when choosing between retailers through discrete choice modelling.

As noted further above in Section 3.3, we interpret the coefficient estimates as representing preferences of both individual shoppers and the households in which they live. In the data, the three variables price, discount and time, account for significant variation in choice of grocery retailer among all the variables entering the model with no interactions. We find that the price variable accounts for the most variation in the data and this variable corresponds to the average basket price a consumer can expect to pay when visiting the grocery store. The variable capturing travel time to the grocery store, accounts for the second most variation in the data out of these three variables, followed by the loyalty scheme discount. All three variables exhibited preference heterogeneity i.e. preferences were not constant for these three variables. When looking at the distribution in tastes (i.e. unobservable variation in preferences) among grocery shoppers, the discount variable displays the most variation with 68% of grocery shoppers favouring a loyalty scheme when choosing between grocery retailers. While the remaining 32% of shoppers prefer not to participate in any loyalty scheme.

Our results show that households have a strong preference for grocery retailers who offer high to very high levels of customer service, product range and product quality. The data revealed that respondents exhibited preference heterogeneity for high levels of product quality, very high levels or product range and very high level of customer service. The remaining qualitative attributes were found to have fixed mean coefficient values among respondents. In assessing the qualitative variables in terms of how they impact consumers' choice of grocery retailer, we note that this data was collected via survey. When answering hypothetical questions, it is likely that some individuals may believe that in a real world context they would actually choose the retailer offering the highest levels of customer service, highest quality of products and a massive range to choose from. In reality however, the majority of consumers are likely to be driven by price and the location of the shop. We are not suggesting that these non-price aspects of the grocery shopping experience, do not matter to consumers. We do however note, that the estimates may be slightly inflated as a result of this effect. As we discussed in the methodological chapter, stated preference analysis has its limitations because in the context of experiments, individuals do not always respond in the same way as they would in a real-world situation.

When analysing our data, we also looked at whether there was any variation in preferences within identifiable groups of consumers (i.e. observable taste variation). The 8 interaction terms found to be significant in model ML4 were:

- Female*VH Range": the effect of gender on preferences for very high range of products at the grocery store;
- "Large Household*VH Range": the effect of living in a large household of four or more individuals, on preferences for very high range of products at the grocery store;
- "No Car*VH Range": the effect of not driving a car to go grocery shopping on preferences for very high product range;
- "< £22,000 HI*Price": the effect of being in the lower household income group earning less than £22,000 annually (after tax) on preferences for grocery basket price;
- "> £45,000 HI*High Service": the effect of being in a higher household income group earning more than £45,000 annually (after tax) on preference for very high levels of service in store;

- "Frequent Online Shop*Price": the effect of regularly purchasing groceries online (i.e. at least once a month) on preferences for the grocery basket price;
- "Frequent Online Shop*Time": the effect of regularly purchasing groceries online (i.e. at least once a month) on preferences for grocery store proximity; and
- "No Loyalty Cards*Discount": the effect of not participating in any loyalty schemes on preferences for the loyalty scheme discount.

In our data, female respondents displayed a lower preference for very high levels of product quality than male respondents. Participants who indicated they lived in a large household displayed a greater preference for very high product range availability than those who indicated they lived in smaller sized households. We also find that those individuals who typically do not drive a car to go grocery shopping, prefer a grocery retailer who offers a bigger product range. Unsurprisingly, we find that households in the lowest income category have a stronger preference for lower average basket prices than households in higher income groups. While those grocery shoppers who live in households in the highest income category, prefer a grocery retailer with high levels of customer service than those who live in households earning lower disposable incomes.

There are also differences in the preferences of individuals who are frequent online shoppers. We find that these individuals are less sensitive to the basket price and the travelling time when choosing their preferred grocery retailer compared to those who are either infrequent online shoppers, or those who never shop online at all. The final interaction we consider, is between loyalty card ownership and the loyalty scheme discount. The results show that while the loyalty scheme coefficient is positive for most individuals, this is not constant for everyone. On the basis of our estimates, respondents who indicated that they do not participate in any loyalty schemes at all, prefer to not have a loyalty scheme at all when choosing between retailers. This further emphasises differences between customer groups in terms of their tastes for loyalty schemes. We keep the above findings in mind throughout the discussion which proceeds below. Next, we consider specific aspects of our results and how they compare to real-world markets.

In Section 3.2, we considered the data collection method we adopted and whether we could expect meaningful results on the basis of this data. In assessing whether our results would be reasonable, we computed implied market shares of the four retailers labelled A, B, C and D, as chosen by the survey respondents. Although they were labelled as A, B, C and D, these four retailer profiles were designed by gathering information on actual retailer characteristics, including prices, namely, Tesco, Sainsbury's, Asda and Waitrose. The market shares we computed were very similar to the actual real world market shares of the four major UK retailers during the same year that the survey responses were collected. This suggests that variation in our data closely follows type of variation in retailer preferences observed in the UK population. In addition we applied population weights to avoid sources of bias affecting the robustness of results as we found a divergence between some of the sample-level statistics versus population-level sociodemographic statistics. The standard errors of estimates have also been clustered at the level of the individual to improve the robustness of results. In consideration of these points, we proceed with the confidence that the empirical results presented in this chapter offer meaningful insights into behaviour of consumers in actual markets.

When designing the attributes to include in the discrete choice experiment, we relied on some of the findings of the Competition Commission's market investigation into the UK groceries market. The CC's Final Report published following the completion of the investigation in 2008, noted that grocery retailers mainly compete for customers through price, quality, product range and service, also known as "PQRS". However, the CC's assessment did not factor in loyalty schemes. Below we consider how our findings sit in the context of the demand estimation procedures performed by the CC using an extensive time-series household dataset. The main findings of the demand estimation were published in the Final Report.

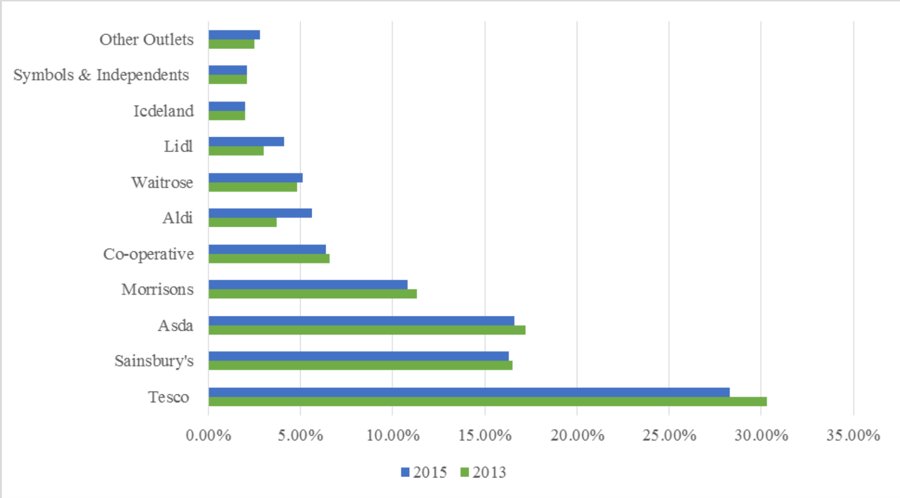
The CC found that preferences for retailers could only be explained by sociodemographic variables to an extent. Instead, using a mixed logit model with random parameters allowed the CC to capture additional aspects of taste heterogeneity. Overall our results are comparable with the CC's findings when considering the effect of store proximity, consumers' preferences for product range and the level of service. The CC's demand model produced a statistically significant and negative effect associated with distance from the store, a positive effects for product availability and levels of service. The CC's report also found evidence of differences in preferences for store proximity and product availability as a result of differences in car ownership and household size. Specifically, the CC's analysis of results found that larger households, and those who own a car, are more likely to travel greater distances to a given grocery store (Competition Commission 2008, Annex 4 p.18). This corresponds to a similar effect we discussed as part of our results where we found a significant relationship between differences in household size and availability of products in the store.

As mentioned above, the CC investigation identified that an important aspect of consumer satisfaction was the presence of a grocery store at close proximity, irrespective of the retailer brand itself. During its investigation, the CC identified that there was substantial gain in consumer satisfaction from having a grocery store within 5 minutes of driving distance. This is also interesting form an exogenous switching cost perspective. This suggests that consumers in the UK groceries sector suffer less from brand related exogenous switching costs than other markets. Our results also show that consumers have a strong preference for having a store at close proximity confirming that proximity remains a key determinant of grocery retailer choice.

In the UK groceries sector, discount stores are also popular which is consistent with the importance consumers place on price. The CC's market investigation concluded in 2008 and since then, discounters Lidl and Aldi experienced growth in market share in the UK. This has been partially at the expense of other leading grocery retailers. The increase in popularity of the discount retailers is consistent with consumers caring a significant amount about the prices they pay. On the other hand, the increase in market shares of Aldi and Lidl, also shows that a lot of consumers do *not* really care about the shopping experience itself where the shopping experience is captured through quality, range and service attributes. As noted above, we also found that average basket price accounted for the most variation in consumers' choice of grocery retailer. In addition, although our data also shows that consumers are likely to care about the quality of products, customer service and product range, these effects are likely overstated for the reasons previously outlined above.

The evolution in the market shares in the UK's groceries market is presented in Figure 3.7 below. The top bar represents August 2015 market shares and the lower bar represents August 2013 market share data respectively. The figure shows a visible growth of market shares of the discounters and a contraction in the market share of the leading retailer Tesco.

Figure 3.7 – UK Grocery Retailer Market Shares (top bar shows August 2015 market share figures, lower bar shows August 2013 market share figures)[69]



The importance of price is consistent with findings of a recent IGD report on drivers of grocery shopper loyalty.[70] Based on the results of IGD's shopper insight survey, convenience (i.e. proximity) and price are by far the biggest drivers of store loyalty. [71] In addition, the IGD report finds that loyalty schemes are also valuable to some shoppers. In the report, 33% of surveyed shoppers stated that a store loyalty card is the main reason for their loyalty to a given grocery retailer while 44% of respondents stated that a loyalty card was either an "extremely important" or "very important" driver of store choice. Thus while loyalty schemes may be very important to some grocery shoppers, this certainly is not the case for all. This also corresponds to our empirical results.

We now apply our empirical findings to the literature reviewed in the first chapter. We focus mainly on the assumptions on consumer preferences and how consumers react when firms adopt repeat purchase discount strategies. In doing so we comment on the likely implications of our findings for the firm strategies and market outcomes suggested by the theoretical models. On the basis of our empirical results, we argue that unlike the typical assumptions entering theoretical models, loyalty schemes do not unilaterally increases costs of switching for consumers. The coefficient estimates we obtained indicate that around a third of consumers and/ or households prefer *not* to receive a repeat purchase discount and would not want to participate in a

[69] Market share data sourced from Kantar Worldpanel, http://www.kantarworldpanel.com/en/grocery-market-share/great-britain.
[70] IGD ShopperVista Report 'Shopper loyalty in 2015', April 2015.
[71] The report is based on a survey of 943 UK grocery shoppers.

loyalty program. For this sub-set of the population, the choice of retailer is independent of the availability of a repeat purchase discount. In fact, our results indicate that a loyalty scheme may actually reduce utility for some individuals. For example, these individuals may not like the idea of their data being collected, stored and analysed by companies. We apply this result to the some of the main papers we discussed in the first chapter on loyalty scheme strategies.

In the literature review chapter we illustrated and solved Lal and Bell's (2003) model of loyalty discounts and exogenously set product promotions. This model is set up differently to the models of loyalty rewarding schemes in the industrial organisation literature. It assumes that some consumers are simply loyal due their proximity to a grocery store following a Hotelling framework. There are promotions which are determined exogenously and the authors consider a scenario when one of the two retailers offers a loyalty scheme on top of the promotional product prices. Consumers who are located in the middle segment of the unit line are cherry-pickers who look for the best deals. Loyalty schemes are shown to *not* affect the behaviour of consumers who are loyal to either one of the two retailers located at the ends of the linear city. These consumers do however, redeem the repeat purchase discount which reduces the retailer's profit. In addition, the model is set up so that only the behaviour of those consumers who cherry pick between stores is affected by the loyalty scheme. This result holds so long as the opportunity cost of shopping around for a better deal is sufficiently compensated by the loyalty discount.

The empirical results presented above are consistent with Lal and Bell's assumption that not all consumers will base their purchasing decisions on loyalty schemes. Of course, the model is relatively simplistic in its form and does not account for the fact that firm's may respond to the consumers' behaviour in future periods. The model does not consider the wider features of a retail market which may also affect consumers' choice of retailer. In addition, our empirical results suggest that not all the consumers would redeem their repeat purchase discount even if they had earned one by shopping at the same retailer. If fewer consumers redeem repeat purchase discounts *and* they continue to choose that retailer, the firm would achieve greater profits than suggested by Lal and Bell's model. The same applies to the Caminal and Claici model which we consider below.

Caminal and Claici's (2007) model assumes that consumers vary in their preferences for variety and are heterogeneous in respect of the brands available.

Caminal and Claici's model also captures the profit incentive driving firms to invest in loyalty schemes when there are few firms in the market. The endogenously created switching costs lock in a segment of consumers enabling the firm to charge higher prices in future. The investment however, reduces profits of the firm and the firm will only offer a repeat purchase discount if the additional profits exceed the cost of the strategy. In the model, consumers are also uncertain about their future preferences and firms are able to discriminate between consumers. A number of model variations are considered, including an extension to multiple periods of competition and where firms discriminate between generations of consumers. This offers a more sophisticated interpretation of consumer preferences than the Lal and Bell model.

In the model, when a firm implements a loyalty scheme, consumers' transport costs increase due the endogenously created switching costs and in certain cases this reduces overall welfare. On the basis of our results we would argue that this increase in transport costs would affect fewer consumers than suggested by the model. This implies that a loyalty scheme would have an impact on future strategies of firms. In this context, with the knowledge that not all consumers redeem repeat purchase discounts, the forward looking firm may for example, decide to offer larger or smaller repeat purchase discounts when maximising pay offs. Further, if firms are able to effectively discriminate between consumers with different switching cost *types* via the loyalty scheme, this would have further implications for firm strategies and market outcomes.

Caminal (2012) also argues that discount strategies may not achieve the most efficient outcome because in practice, a future price commitment achieves the most efficient outcome in terms of welfare. Without a commitment in place, a firm can raise prices unilaterally to compensate for the reduction in profit from offering customers a repeat purchase discount. In this context, some consumers will find themselves in a prisoner's dilemma, where they are equally better off not participating in the scheme as they end up facing higher prices in future periods. More generally, Caminal (2012) notes that the models of endogenous switching costs, namely the models above as well as the model of Caminal and Matutes (1990), are set up in a way that "*LRs allow firms to retain previous customers, even when rival firms offer goods or services that better match their current preferences. As a result, LRs are welfare reducing because they* *cause a mismatch in the allocation of consumers.* […] *it is unclear whether LRs tend to relax or exacerbate price competition"*[72]

In light of our empirical results, we argue that the reduction in welfare suggested by a broad range of models of endogenous switching costs would be less pronounced. As noted above, the results suggest that when choosing between retailers, a third of consumers prefer not to have and/ or participate in a loyalty scheme. This suggests that some consumers may even be deterred by the scheme. Instead, choose the retailer corresponding to their current preferences, even if it is the retailer who offers the loyalty scheme. In real world markets, consumers are not forced into redeeming loyalty points in the form of lump sum coupons or other rewards. It is therefore unrealistic to assume that when a firm implements a loyalty rewarding scheme, this unilaterally increases artificial switching costs in consumers. We note that variation in the way consumers incur endogenous switching costs, may have ambiguous welfare effects depending on the model design, namely whether the firms can observe this aspect of behaviour and how they would react in response. Considering the above, we argue that the strategy is likely to have a weaker effect on price competition than suggested by the literature. While the extent that consumers are heterogeneous in their switching costs is only one of many assumptions entering a model, it would be of interest to understand how this assumption would affect the outcomes suggested in the theory.

We now consider the implications of our findings for competition policy. Let us first consider the UK groceries market. In this sector, firms compete over various aspects of the retail offer to attract consumers to their stores and invest heavily in branding and advertising. This is supported by our DCE results. Since the CC's market investigation which concluded in 2008, the main players, in particular Tesco who previously dominated the market, experienced fluctuating performances as a result of changing market conditions, poor investments and accountancy related scandals. More generally, we noted above that retailers face competitive pressures from shifting shopping patterns towards convenience and pressures from newer expanding entrants such as discount retailers Aldi and Lidl.

Considering these wider aspects of the UK groceries market, no one retailer is likely to enjoy significant market power. Therefore on the basis of our results, loyalty

[72] Caminal, R., 'The Design and Efficiency of Loyalty Rewards', *Journal of Economics and Management Strategy*, Vol. 21, Nr. 2, 2012, p. 340

schemes are unlikely to be an anticompetitive device in this market. Instead, loyalty schemes in this market represent one of *many* aspects of the competitive process in a mature market. In reality, a loyalty scheme, in this particular market, may have a locking-in effect on certain *groups* of consumers. For example those indicating they participate in no loyalty schemes at all are found to prefer not receiving a loyalty scheme discount. This is consistent with Section 1.5 of the first chapter. In the context of the retail energy and retail banking market investigations, the surveys commissioned by the CMA suggested that some groups of consumers were likely to face different costs of switching and were switching at different frequency, with some consumers not switching at all. These findings also suggest that consumers are heterogeneous in their costs of switching in markets with exogenous switching costs. For example in the context of retail banking, younger consumers are more tech savvy and can be expected to be more engaged and willing to switch (Banking MI, para 5.165). On balance, loyalty schemes are unlikely to be problematic from a competition policy perspective in dynamic markets for nondurables assuming that consumers face minimal exogenous switching costs and firms continually invest in maintaining market share, for example, through lower prices and/ or better quality.

Given the significant weight placed on consumer outcomes by EU competition authorities, understanding consumer behaviour is an important first step in a competition related investigation. The favourable statistical properties and insights offered by the discrete choice experiment outlined in this thesis, could also be of value to competition authorities or researchers in industrial organization. DCE can help investigate the effects of a business strategy on consumer behaviour or identify how preferences are distributed in the population of interest. In addition, the approach outlined in this thesis can help assist in policy design. For example, if the competition authority is seeking to have a better understanding of consumer preferences in a market associated with consumer switching costs. The approach may help support a broader analysis to identify the relevant policies to reduce switching costs and increase switching by consumers for example.

Compared to a discrete choice experiment, theoretical models may not capture the wider aspects of the retail offer. In particular, such models are less able to accommodate diverse consumer preferences and the differentiated characteristics of sellers which prevail in real-world markets. However, in evaluating our approach we also note that in the context of DCEs, consumers may not respond in the same way as they do in real world markets. In addition, experiments more generally are imperfect due to the complex nature of competitive interactions in markets namely, changes to strategies in response to changes in others' strategies. Thus both approaches suffer from drawbacks. We therefore argue that when analysing the effects of business strategies which rely heavily on assumptions on consumer choices, a combination of theoretical and experimental evidence may be optimal.

An advantage of performing a DCE compared to the above model, is that we were able to ass loyalty schemes in the context of price and non-price factors which matter to consumers. On balance, DCEs enable the researcher to address very specific questions on drivers of consumer choice by mimicking real-world markets through a series of survey questions or controlled lab experiments. Importantly, the approach accommodates all types of behavioural patterns, including non-utility maximising behaviour. Thus, empirical evidence can help determine realistic assumptions to enter a theoretical model which can then be applied to achieve broader analysis of business strategies and its implications for policy which consider both the demand and supply side of the market.

In this context, theoretical modelling is essential in explaining broader dynamics of markets to help us understand the underlying rationale and incentives for the pricing and discounting strategies adopted by firms, including loyalty schemes. However, this approach focuses on the firm rather than the consumer. It is therefore of essence to accurately introduce consumer behaviour into such models and their responses to various firm strategies. Thus, in the assessment of strategic behaviour of firms which prevail in real-world markets, we favour the approach which combines insights from theoretical and empirical modelling. This is more likely to overcome some of the trade-offs associated with theoretical and experimental approaches and is therefore more likely to offer the greatest insights on different aspects of a market. For example, looking back at Shi's (2012) model of exogenous *and* endogenous switching costs. Our modelling exercise did not include brand names as we deliberately designed an unlabelled choice experiment which is discussed at length in the methodological chapter. However, we note that it would be of interest to empirically assess whether the brand attachment effect, or other exogenous switching costs would be greater than that of a loyalty scheme when considering actual consumer behaviour. Let us look at some further areas for future research.

Loyalty cards typically enable firms to collect vast amounts of data to improve their knowledge of customers' behaviour and preferences. In the United States and United Kingdom, loyalty schemes have achieved popularity among both firms and their customers. In the US, the second largest chemist CVS has a loyalty scheme with 69 million subscribers.[73] The loyalty scheme strategy aims to maintain existing customers using regular price discounts and targeted coupons based on purchase history.[74] Similarly to the US experience, a report by YouGov (2013) on British shoppers, identified that 76% typically carry one to five loyalty cards in their wallets. The report also finds that 32% of shoppers are willing to exchange further personal information in return for extra loyalty points. In this context, loyalty schemes offer perceived benefits to both firms and consumers.

On the basis of our findings, we argue that loyalty schemes are likely to continue to play a role as a differentiation mechanism for firms. For example, Tesco have developed novel mobile payments via a new app. This allows customers to quickly pay for their shopping and collect points all in one go as their payment details are linked to the app. The payment facility could easily be expanded to a digital wallet loyalty scheme allowing customers to collect and spend points across different channels. This is already a possibility with some credit cards.[75] In this context, in future, firms will still likely be able to rely on loyalty schemes as mechanisms for differentiation. Thus future research should consider the competitive implications of loyalty schemes which affect multiple markets instead of only focusing on one market or one type of product.

In addition to the above strategy, firms may also use loyalty schemes to collect rich customer data. Lal and Bell's model and our empirical results suggest that a loyalty scheme's profitability depends on the firm's ability to identify and target the consumers whose behaviour is affected by the loyalty discount. Firms actively use loyalty reward programs to help them identify consumers using the data that have accumulated. This facilitates targeted marketing and product discounts. The same can be said about firms who operate online and target consumers with adds on the basis of their search history.

---

[73] 'Retailing: Spies in your Wallet', The Economist, November 5th 2011 (available online http://www.economist.com/node/21536604)

[74] *Ibid.*

[75] Research by IGD has already identified existence of "smart" loyalty cards that have been introduced in the United States. See: IGD, 'What impact do loyalty schemes have on store choice?' 15th July 2013, http://www.igd.com/Research/Shopper-Insight/shopper-outlook/15151/What-impact-do-loyalty-schemes-have-on-store-choice/

Therefore firms are likely to have the incentive to invest in scheme effectiveness to target consumers whose behaviour can be impacted through strategic behaviour. This also has implications for competition and a firms' market power. We understand that firms can gain a competitive data advantage over rivals where the *"data's competitive significance (and value) arise in part from the ability of firms to exclude others from access and analysing it as quickly*."[76] There is therefore a growing interest in the role of big data in competition policy. This is in the context of the likely implications of firms' access to proprietary customer data used to inform basic internal company workings as well as strategic business decisions.[77] This is consistent with Tesco's Clubcard loyalty scheme being central to its business which we explained in the first chapter in Section 2.2.

An argument previously outlined in an article in The Economist stated that retailers' investment into loyalty cards is not intended to induce customer loyalty but rather to collect their data.[78] Tesco was the first retailer to implement a loyalty scheme strategy in the UK groceries market and has been collecting and analysing data on customers in this way ever since. The company, Dunnhumby, which helped Tesco establish the Clubcard in the early 1990s, is wholly owned by Tesco. Dunnhumby is also a leading firm in customer data science for retailers and brands.[79] While this highlights Tesco's data advantage, *if*, or to what *extent*, Tesco's market share can be explained by this aspect of its business is another question altogether. Thus while this was not the focus of our thesis, it offers a fruitful direction for future research. In this context, it would be of interest to model and assess the competitive data advantage conferred to firms offering loyalty programs. Additionally, future research should consider the impact of companies' acquisition and analysis of vast customer data on their strategic decisions and resultant market power, both in online as well as brick and mortar channels respectively.

3.7 Conclusion

This chapter presented the results of the discrete choice experiment on the UK groceries market developed as part of the methodological chapter. In Section 3.2 we

[76] Stucke, M., Allen Grunes, A., *Big Data and Competition Policy,* Oxford University Press, paragraph 4.26, 2016.
[77] *Ibid.* Part I, Section 4, *The Competitive Significance of Big Data.*
[78] 'Retailing: Spies in your Wallet', The Economist, November 5th 2011 (available online http://www.economist.com/node/21536604)
[79] https://www.tescoplc.com/about-us/our-businesses/dunnhumby/about-the-business/

explained the processes undertaken to evaluate the quality of the survey data and we also tested the data for sources of bias through a comparative assessment on the basis of population-level statistics in Section 3.3. We were then able to control for sources of sample bias through the application of frequency weights. In sections 3.4 and 3.5 we presented a number of model outputs based on different specifications of the conditional logit and mixed logit models. Section 3.6 applied the empirical results to the theoretical literature presented in the first chapter and discussed the implications of our results to competition policy. We critically assessed our approach with respect to theoretical modelling and concluded with a discussion on directions for further research on loyalty scheme strategies.

As part our results, we found that individuals have a strong preference for non-price aspects of the grocery store offering, namely the product quality, range and level of customer service. For example, 97% were found to prefer high levels of product quality. We explained the possibility that these effects may be overstated because the results were obtained using *stated preference* instead of *revealed preference* data. As such, we concluded that on balance, the majority of households prefer lower prices over and above other retailer features. Further, we found that households are heterogeneous in their preference for a number of price and nonprice grocery retailer attributes. Notably, of all the variables entering the specification, preferences were the most varied for the loyalty reward scheme with a third of consumers preferring not to have a loyalty scheme when choosing between retailers.

In considering the outcomes suggested in the theory, we recalled that firms have the profit incentive to offer strategic discounts as this increases artificial switching costs in consumers. On that basis, firms can increase prices in future periods if no commitments are in place. In this type of set up, loyalty rewarding schemes can be shown to impact competition in different ways, either having a softening or intensifying effect. Outcomes tend to depend on the market structure, the type of price commitment in place and number of periods entering the theoretical model. Our results indicate that because consumers are heterogeneous in their switching costs, the effects of repeat purchase discounts do not have a consistent effect across the population of grocery shoppers in the UK.

On the basis of our results, we conclude that loyalty schemes do not create artificial switching costs for all consumers, at least not to the same degree. In turn, the effects of the strategy are likely to be weaker and produce a milder impact on price competition that suggested in the literature. It is therefore unrealistic to assume that when a firm implements a loyalty rewarding scheme that this will unilaterally increase artificial switching costs in all consumers. This also suggests that firms may be less incentivised to engage in *harvesting* of consumers and will choose to *invest* in market share instead. The investment incentive will be strong if most consumers are active switchers in a non-durable goods market with low brand effect related switching costs. While this has not been tested as part of this thesis, lack of brand attachment in grocery retail may also explain why competition in this market is strongly driven by price. We also note that in practice, grocery retailers are unable to price discriminate in a material way between different groups of consumers with varying sensitivities to loyalty rewarding schemes. The evidence and arguments outlined in this chapter suggest that retailers are unable to rely on loyalty schemes alone to retain and build their market share. Instead, it is more likely that such firms must rely on other levers of competition to attract different *types* of consumers, namely by choosing lucrative geographic locations, offering better service, higher quality products or lower average prices.

Appendix

Table A.2.1 – Raw basket price data using ONS CPI list of 139 food and household items most frequently purchased by UK households. The table shows the prices and item descriptions for the cheapest own brand product at Tesco, Sainsbury, Asda and Waitrose. The data was collected using the grocery retailer online websites during January 2014.

		Tesco	Sainsbury's	Asda	Waitrose
	Total CPI Basket Items Price	£216.64	£234.34	£201.32	£305.57
Item	**Item Description**		**Price**		
A	**Bread and Cereals**				
1	*Cereal bars (cheapest option 150 gr)*	0.92	0.99	0.83	1.92
2	Chocolate wafers	**N/A**	**N/A**	**N/A**	**N/A**
3	Corn based snacks (Cheese Puffs 100 gr)	0.66	0.64	0.66	0.66
4	*Cornflakes (500 gr)*	1.29	1.29	1.28	1.29
5	*Cream Crackers (300 gr)*	0.45	0.36	0.50	0.36
6	*Crusty Bread rolls (4 Fresh from Bakery)*	0.65	0.70	0.70	1.40
7	*Flour (Plain 1.5 kg)*	0.60	0.60	0.60	1.11
8	*Frozen pizza (Cheese and Tomato 250-300 gr)*	0.60	0.60	0.60	1.84
9	*Fruit pies (Fresh Bramley apple pie)*	1.00	1.10	1.00	2.50
10	*Garlic bread (twin pack 420 gr)*	1.50	1.50	1.20	1.50
11	*Hot oat cereal (porridge 1 kg)*	1.20	1.20	0.85	1.10
12	*Jam Doughnuts (5 pcs)*	0.65	0.65	0.65	0.69
13	*Large white loaf (Fresh Sandwich Loaf 800 gr)*	1.30	1.25	1.30	1.28
14	*Large wholemeal loaf (medium sliced 800 gr)*	0.50	0.50	0.47	0.80
15	*Long Grain Rice (1 kg)*	0.40	0.44	0.40	1.39
16	*Pack of individual cakes (Chocolate chip muffins 4 pk)*	1.50	1.50	1.00	1.50
17	*Fresh Pasta (Penne 500 gr)*	1.72	1.60	1.43	1.70
18	*Pasta (Dry Spaghetti 500 gr)*	0.39	0.39	0.39	0.95
19	*Sponge cakes (Victoria Sponge Cake)*	2.20	2.30	2.48	2.29
20	*Various selected biscuits (Custard Creams 400gr)*	0.40	0.40	0.31	0.90
B	**Meat**				
	Beef				
21	*Beef mince (500gr)*	1.56	1.46	1.46	3.19
22	*Braising steak (cheapest available 500 gr)*	4.00	4.00	4.00	5.00
23	Frozen burgers	**N/A**	**N/A**	**N/A**	**N/A**
24	*Rump Steak (500 gr)*	5.00	6.32	5.50	8.75
	Topside (Joint 500 gr)	5.00	5.00	5.39	5.65
25	**Lamb**				
26	*Leg of Lamb (1 KG)*	10.99	10.99	6.00	12.99
27	*Loin chops (500 gr)*	4.00	5.13	4.50	8.50
28	*Shoulder (half shoulder joint 1 KG)*	7.00	7.00	5.50	7.99
	Pork				
29	*Loin chops (500 gr)*	2.43	2.43	2.00	3.73
30	*Bacon (300 gr)*	1.64	1.64	1.70	2.85
31	*Gammon (Steaks 500 gr)*	2.60	4.69	2.25	5.00

	Chicken				
32	*Chicken Breasts (500 gr)*	6.00	6.50	4.85	6.89
33	*Frozen Chicken Nuggets (300 gr)*	1.00	1.64	1.04	3.00
34	Fresh/chilled whole chicken (1.5 kg)	3.72	3.75	3.72	4.92
35	Frozen chicken breasts (1kg)	4.75	6.49	3.99	8.80
36	Rotisserie cooked hot whole chicken	**N/A**	**N/A**	**N/A**	**N/A**
	Other Meats				
37	*Steak Pie (550 gr)*	3.50	3.65	3.38	3.59
38	*Pork Sausages (8pcs)*	0.61	0.65	0.44	1.39
39	*Cooked meats – eg ham (roast turkey breasts 200 gr)*	2.70	3.34	4.16	5.58
40	*Fresh turkey steaks (4 pk 500 gr)*	4.54	4.50	3.69	5.00
41	*Canned meats (corned beef 300 gr)*	2.02	2.01	1.36	3.12
42	*Frozen Chicken Nuggets (300 gr)*	0.68	1.17	0.68	3.00
43	*Chicken kievs (2 pk)*	1.39	1.98	2.18	2.27
44	*Oven-ready joint (pork belly 1kg)*	5.00	5.33	5.30	6.39
C	**Fish**				
45	*Frozen Breaded Cod Fillets (500 gr)*	2.20	3.20	2.00	3.00
46	*Fresh white fish fillets (Pre-Packed Cod Fillets 250 gr)*	2.98	2.98	3.25	3.66
47	*Fresh salmon fillets (pre-packed 300 gr)*	2.69	2.69	2.69	4.08
48	Canned tuna (185 gr)	0.75	0.75	0.75	1.71
49	*Fish fingers (Frozen Cod 300 gr)*	1.75	1.91	1.41	1.91
50	*Frozen prawns (Cooked and peeled King prawns 250 gr)*	3.25	2.99	3.58	3.82
D	**Milk, Cheese and Eggs**				
51	*Cheese spread (Soft cheese 200 gr)*	0.49	0.60	0.49	0.80
52	*Chilled pot dessert (Chocolate Mousse Pack 6 x 62.5 = 375 gr)*	0.33	0.33	0.33	0.90
53	*Edam (wedge 310 gr)*	2.00	2.16	2.00	2.55
54	*English Medium Cheddar (300gr)*	2.26	2.36	2.25	2.49
55	*Fresh cream (Single 300 ml)*	1.05	0.95	0.95	0.95
56	Fromage frais (low fat 500 gr)	1.00	1.00	1.00	1.19
57	*Medium Free Range Eggs (12 pcs)*	2.65	2.65	1.98	2.92
58	Milk (6 pints half fat)	1.89	1.89	1.48	1.89
59	*Other regional cheeses (Mozzarella 125 gr)*	0.44	0.65	0.44	0.95
60	*Parmesan (200 gr)*	3.25	3.29	3.20	3.72
61	Powdered baby formula	**N/A**	**N/A**	**N/A**	**N/A**
62	Pro-biotic drink	**N/A**	**N/A**	**N/A**	**N/A**
63	Soft continental cheese (French Brie 200 gr)	1.09	1.09	1.00	1.67
64	Yoghurt (Natural Low-Fat 500 gr)	0.49	0.65	0.49	1.00
E	**Oils and Fats**				
65	*Butter Salted (250 gr)*	1.19	1.20	0.98	1.20
66	*Margarine/low fat spread (Olive spread 500 gr)*	1.39	1.50	1.39	1.50
67	*Olive oil (1 lt)*	3.80	3.69	3.48	3.69
F	**Fruit**				
68	*Pineapple (1 pc)*	1.00	1.00	0.80	1.69
69	*Avocado (ready to eat twin pack)*	2.00	2.00	1.75	1.99
70	*Bananas (Loose 1 kg)*	0.79	0.79	0.68	0.79
71	*Cooking apples (1 kg)*	1.95	1.99	1.95	1.99

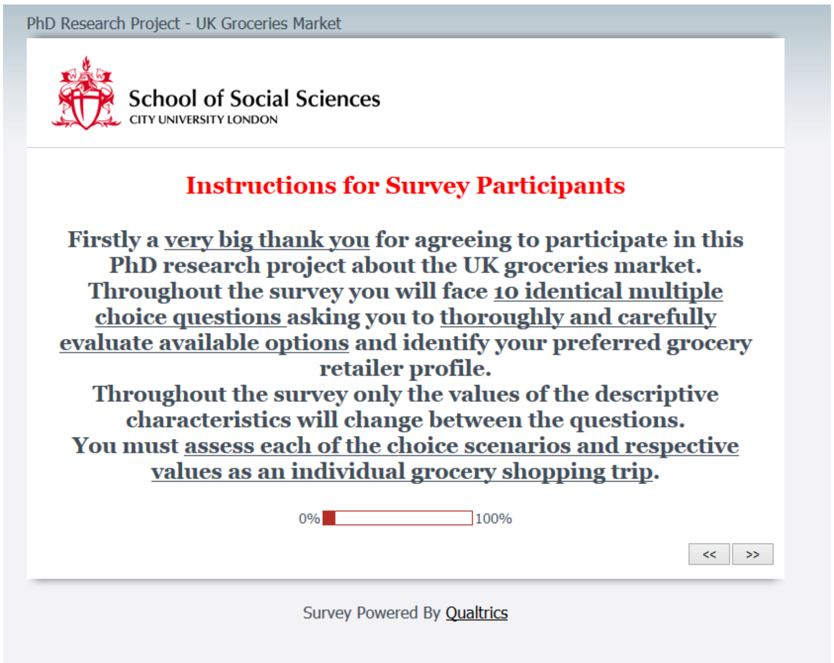
72	*Dessert apples (Gala Bag of 6)*	1.70	1.99	1.50	1.99
73	*Dried fruit (Dried Mango 100 gr)*	1.50	1.75	1.47	1.93
74	*Grapefruit (Red x 3)*	1.50	1.41	1.08	1.44
75	*Green Seedless Grapes Pack (500 gr)*	2.25	2.50	2.00	2.75
76	*Kiwi fruit (6 pk)*	1.00	1.25	1.00	1.25
77	*Oranges (Loose 5 pcs)*	1.50	1.75	1.50	1.50
78	*Organic fruit (Lemons 4pk)*	1.33	1.50	1.87	1.99
79	*Peaches/nectarines (Punet of 4)*	1.50	2.50	1.75	3.00
80	*Pears (Ripe 4 pk)*	1.75	1.80	1.75	2.50
81	*Plums (loose 500 gr)*	2.50	2.50	2.50	2.50
82	*Salted/roasted peanuts (200 gr)*	0.55	0.72	0.48	1.08
83	*Small oranges (bag of 5)*	0.59	0.93	0.48	1.99
84	*Strawberries (300 gr)*	2.31	1.80	1.32	3.80
85	*Various canned fruits (pineapple pieces 500 gr)*	0.30	0.71	0.28	0.92
G	**Vegetables**				
86	*Broccoli (1 unit 335 gr)*	1.00	1.00	0.50	1.69
87	*Onions (Red 3 pk)*	1.00	1.00	0.90	1.25
88	*Baking Potatoes (Bag 2.5 KG)*	2.40	2.60	2.20	2.75
89	*Cabbage (Savoy 1 pc)*	0.80	0.80	0.80	0.80
90	*Canned baked beans (4x420 gr)*	1.60	1.60	1.27	1.68
91	*Canned sweet corn (325 gr)*	0.62	0.69	0.59	0.69
92	*Canned tomatoes (4x400 gr)*	2.19	2.49	1.56	2.49
93	*Carrots (loose 1 kg)*	0.90	0.90	0.90	0.90
94	*Cauliflower*	0.89	1.00	0.50	1.00
95	*Courgettes (loose 1 kg)*	1.90	2.00	1.62	2.00
96	*Crisps – single and multi-packs (sea salt 150 gr)*	1.39	1.50	1.00	1.50
97	*Cucumbers (1 pc)*	0.80	0.80	0.50	0.90
98	*Frozen chips (1.5 kg)*	1.00	1.00	0.93	1.41
99	*Frozen Garden Peas (1kg)*	1.60	1.76	1.60	1.77
100	*Lettuce (Round 1 pc)*	0.57	0.60	0.57	0.60
101	*Mushrooms (White Closed Cup 400 gr)*	0.97	0.97	1.07	1.30
102	*Organic vegetables (Organic Leeks 400 gr)*	2.00	2.00	2.00	2.00
103	*Peppers (Mixed Bag 600 gr)*	1.34	1.51	1.52	1.75
104	*Pre-packed salad (Leafy Rocket Salad 100 gr)*	1.39	1.88	1.11	1.66
105	*Tomatoes (500 gr)*	0.95	0.90	0.89	1.69
106	*Vegetable pickle (Onion 440 gr)*	0.30	0.31	0.30	0.99
107	*Vegetarian burger/grills*	**N/A**	**N/A**	**N/A**	**N/A**
H	**Sugar, Jam, Honey, Syrups, Chocolate and Confectionery**				
108	*Chocolates (Milk chocolate bar 200 gr)*	0.60	0.66	0.60	1.62
109	*Gum*	**N/A**	**N/A**	**N/A**	**N/A**
110	*Ice cream (Vanilla 2 litres)*	1.00	0.89	0.89	1.50
111	*Mints (Assortment 200 gr)*	0.89	0.78	0.80	1.00
112	*Sugar*	**N/A**	**N/A**	**N/A**	**N/A**
113	*Various jams (Strawberry Jam 454 gr)*	0.29	0.29	0.29	0.80
114	*Various selected popular brands of sweets*	**N/A**	**N/A**	**N/A**	**N/A**
I	**Food Products (not elsewhere classified)**				
115	*Mayonnaise (500 ml)*	0.45	0.50	0.45	1.09

116	Ready cooked meals (Fresh Cottage Pie 450 gr)	2.30	2.30	2.25	2.69
117	*Soup (Leek and Potato 600 gr)*	1.00	1.70	1.00	1.99
118	*Tomato Sauce (Squeezy Ketchup 500 gr)*	0.92	1.03	0.87	1.03
J	**Non-alcoholic Beverages**				
	Coffee, Tea and Cocoa				
119	*Tea Bags (80 bags)*	0.27	0.35	0.35	1.50
120	*Instant Coffee (Rich Roast 100 gr)*	1.50	1.50	1.84	1.50
121	Ground Coffee (227 gr)	1.69	1.69	2.28	2.29
	Mineral Waters, Soft Drinks and Juices				
122	*Cola (2 litres)*	0.17	0.20	0.17	0.95
123	Energy drinks	**N/A**	**N/A**	**N/A**	**N/A**
124	*Fruit drink (Cranberry Juice 1 litre)*	1.00	1.00	1.00	1.80
125	*Fruit smoothie (Tropical Fruit 1 litre)*	1.33	2.00	1.33	2.39
126	*Lemonade (2 litres)*	0.17	0.20	0.18	0.69
127	*Mineral water (Sparkling 4x2 litres)*	1.50	1.65	1.50	1.65
128	*Squash (Orange Double Strength 1.5 litres)*	1.50	1.59	1.49	2.24
129	*Various fizzy drinks (Ginger Ale 1 litre)*	0.61	0.51	0.45	0.50
130	*Various pure fruit juices (Orange Fresh 1 litre)*	1.20	1.20	1.00	1.20
K	**Goods and Services for Household Maintenance**				
	Non-Durable Household Goods				
131	*Bin liners (Standard Tie Top Refuse 20 Pack)*	2.50	2.50	2.50	2.66
132	*Bleach (Thick Citrus 750 ml)*	0.79	0.87	1.00	0.87
133	*Washing powder (3 kg Bio Powder)*	2.80	2.88	2.52	6.45
134	*Washing-up liquid (500 ml)*	0.33	0.40	0.33	0.89
135	*Aluminium foil (20 m)*	1.13	0.68	0.72	2.98
136	*Dishwasher tablets (30 pk)*	1.60	1.80	1.67	3.70
137	*Fabric conditioner (2 l)*	0.90	1.20	0.90	2.20
138	*Household cleaner cream/liquid (all purpose liquid 1l)*	0.33	0.33	0.33	1.52
139	*Kitchen roll (2 rolls)*	1.25	1.25	1.25	1.58

A.2.2 – Ngene pilot design syntax

```
;alts = T, S, A, W
;rows = 10
;eff = (mnl,d)
;alg = mfederov
;require:
T.P= [56.07] and S.P= [60.65] and A.P= [52.60] and W.P= [80.59],
T.Disc= [0] or T.Disc= [29.16] or T.Disc= [58.31] or T.Disc= [116.62],
S.Disc= [0] or S.Disc= [31.54] or S.Disc= [63.08]
;reject:
W.Qual=0,
A.Qual=3
;model:
U(T) = b1[-.01]*P[52.60, 56.07, 60.65, 80.59] + b2[0.0001]*Disc[0,
29.16, 58.31, 31.54, 63.08, 116.62]
+ b3[-.001]*Time[5, 8, 12, 17] + b4.effects[-.0002|-
.0001|.0001]*Qual[0,1,2,3]
+ b5.effects[-.0002|-.0001|.0001]*Ran[0,1,2,3] + b6.effects[-.0002|-
.0001|.0001]*Serv[0,1,2,3]
+ i1[.0001]*Disc*Time + i2[-.0001]*P*Time + i3[.0001]*Disc*P +
i4[.0001]*P*Qual.effects[2]+ i5[.0002]*P*Qual.effects[3]
+ i6[-.0002]*P*Qual.effects[0] + i7[.0001]*P*Serv.effects[2] +
i8[.0002]*P*Serv.effects[3]
+ i9[-.0002]*P*Serv.effects[0]
/
U(S) = b1*P + b2*Disc + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(A) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(W) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
$
```

A.2.3 – Screen shot of pilot survey instructions for participants



A.2.4 – Screenshot of a sample question from the pilot survey

PhD Research Project - UK Groceries Market

School of Social Sciences
CITY UNIVERSITY LONDON

Based on the following characteristics and subject to your current income; please indicate below your preferred retailer profile where you would choose to shop most frequently:				
Grocery Retailer	**A**	**B**	**C**	**D**
Average Weekly Basket Price of Own-Brand Items that are Frequently Purchased by UK Households *(Based on ONS data for 2 person household average weekly spend on food, non-alcoholic drink & basic household items such as foil)*	£56.00	£61.00	£53.00	£81.00
Annual Loyalty Discount Based on the Above Average Basket Price *(The discount value is increasing with total expenditure on all retailer products. Discount coupon can be spent in store, on numerous leisure activities, holidays etc.)*	£117.00	£63.00	£0.00	£0.00
Drive Time to the Store in Minutes	5	8	17	12
Quality of Products in Store *(Based on the standard own-brand items ONLY; excludes high-end items and branded products)*	Low	High	Medium	Very High
Range *(The extent of product variety in store in addition to the actual product categories such as groceries, lottery, electronics, clothing, make-up, banking, etc.)*	Medium	High	Very High	Medium
Standard of Service *(Overall friendliness & helpfulness of staff, type of returns policy, cleanliness of the store, availability of parking spaces, and general in-store experience.)*	Medium	Medium	Medium	Low

- A
- B
- C
- D

0% 100%

<< >>

Survey Powered By Qualtrics

A.2.5 - Pilot survey results before clustering standard errors at individual level

	Model 1		Model 2	
Variable	**Coef.**	**|*z*-stat|**	**Coef.**	**|*z*-stat|**
Price	-0.036*	3.67	0.012	0.11
	(0.01)		(0.111)	
Discount	0.01*	3.47	-0.105	1.63
	(0.003)		(0.065)	
Time	-0.081*	3.62	-0.234	0.8
	(0.022)		(0.293)	
Medium Quality	2.39*	4.35	18.313	1.34
	(0.549)		(13.68)	
High Quality	2.73*	5.25	20.432	1.39
	(0.52)		(14.685)	
Very High Quality	2.938*	5.46	22.92	1.61
	(0.538)		(14.229)	
Medium Range	0.481*	1.84	0.322	0.12
	(0.262)		(2.662)	
High Range	0.576	1.63	-0.820	0.32
	(0.354)		(2.6)	
Very High Range	0.772*	2.71	3.203	1.19
	(0.285)		(2.683)	
Medium Service	0.744*	0.066	0.56	1.44
	(0.271)		(0.388)	
High Service	0.814*	0.103	0.793	1.50
	(0.308)		(0.0529)	
Very High Service	1.101*	0.007	1.521*	4.12
	(0.336)		(0.369)	
Price*Discount			0.002	1.79
			(0.001)	
Price*Time			0.003	0.52
			(0.293)	
Time*Discount			-0.001	1
			(0.001)	
Price*Low Quality			0.293	1.22
			(0.241)	
Price*High Quality			0.293	0.54
			(0.063)	
Price*Very High Quality			-0.067	1.07
			(0.062)	
Price*Low Service			-0.0005	0.01
			(0.044)	
Price*High Service			0.014	0.31
			(0.458)	
Price*Very High Service			-0.051	0.97
			(0.052)	
Log-likelihood	-270.565		-262.059	
Nr. Respondents	26		26	
Nr. Observations	1040		1040	

Notes: Standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

A.2.6 – Ngene main design syntax: model averaging approach with Bayesian approximation

```
Design
;alts(m1) = T, S, A, W
;alts(m2) = T, S, A, W
;alts(m3) = T, S, A, W
;rows = 11
;eff = 2*m1(mnl,d,mean) + 1.5*m2(mnl,d,mean) + m3(mnl,d,mean)
;alg = mfederov
;bdraws= Halton(40000)
;require:
T.P= [56] and S.P= [61] and A.P= [53] and W.P= [81],
T.Disc= [0] or T.Disc= [29] or T.Disc= [58] or T.Disc= [117],
S.Disc= [0] or S.Disc= [32] or S.Disc= [63]
;reject:
W.Qual=0,
A.Qual=3
;model(m1):
U(T) = b1[(n,-.04, .05)]*P[53, 56, 61, 81] + b2[(n,0.01,.02)]*Disc[0,
29, 58, 32, 63, 117] + b3[(n,-.08,.01)]*Time[5, 8, 12, 17]
+ b4.dummy[(n,2.39,2.7)|(n,2.73,2.8)|(n,2.94,2.7)]*Qual[1,2,3,0] +
b5.dummy[(n,.74,1.38)|(n,.81,1.57)|(n,1.1,1.71)]*Ran[1,2,3,0]
+ b6.dummy[(n,.48,1.34)|(n,.58,1.81)|(n,.77,1.71)]*Serv[1,2,3,0]
/
U(S) = b1*P + b2*Disc + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(A) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(W) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv

;model(m2):
U(T) = b1[(n,-.04, .05)]*P[53, 56, 61, 81] + b2[(n,0.01,.02)]*Disc[0,
29, 58, 32, 63, 117] + b3[(n,-.08,.01)]*Time[5, 8, 12, 17]
+ b4.dummy[(n,2.39,2.7)|(n,2.73,2.8)|(n,2.94,2.7)]*Qual[1,2,3,0] +
b5.dummy[(n,.74,1.38)|(n,.81,1.57)|(n,1.1,1.71)]*Ran[1,2,3,0]
+ b6.dummy[(n,.48,1.34)|(n,.58,1.81)|(n,.77,1.71)]*Serv[1,2,3,0]
+ i1[-.001]*Disc*Time + i2[.001]*P*Time + i3[.001]*Disc*P + i4[-
.001]*P*Qual.dummy[2]+ i5[-.001]*P*Qual.dummy[3] +
i6[.001]*P*Qual.dummy[0] + i7[.001]*P*Serv.dummy[2]
+ i8[-.001]*P*Serv.dummy[3] + i9[-.001]*P*Serv.dummy[0]
/
U(S) = b1*P + b2*Disc + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(A) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(W) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv

;model(m3):
U(T) = b1[(n,-.04, .05)]*P[53, 56, 61, 81] + b2[(n,0.01,.02)]*Disc[0,
29, 58, 32, 63, 117] + b3[(n,-.08,.01)]*Time[5, 8, 12, 17]
+ b4.effects[(n,2.39,2.7)|(n,2.73,2.8)|(n,2.94,2.7)]*Qual[1,2,3,0] +
b5.effects[(n,.74,1.38)|(n,.81,1.57)|(n,1.1,1.71)]*Ran[1,2,3,0]
+ b6.effects[(n,.48,1.34)|(n,.58,1.81)|(n,.77,1.71)]*Serv[1,2,3,0]
/
U(S) = b1*P + b2*Disc + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(A) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(W) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
$
```

A.2.7 – Ngene model evaluation syntax for the mixed logit specification

```
Design
;alts(m1) = T, S, A, W
;rows = 11
;eff = m1(rppanel,d)
;rdraws= halton(1000)
;rep = 1000
;eval = 6.ngd
;model(m1):
U(T) = b1[n,-.04, .05]*P[53, 56, 61, 81] + b2[n,0.01,.02]*Disc[0, 29,
58, 32, 63, 117] + b3[n,-.08,.01]*Time[5, 8, 12, 17]
+ b4.dummy[n,2.39,2.7|n,2.73,2.8)|n,2.94,2.7]*Qual[1,2,3,0] +
b5.dummy[n,.74,1.38|n,.81,1.57|n,1.1,1.71]*Ran[1,2,3,0]
+ b6.dummy[n,.48,1.34|n,.58,1.81|n,.77,1.71]*Serv[1,2,3,0]
/
U(S) = b1*P + b2*Disc + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(A) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
/
U(W) = b1*P + b3*Time + b4*Qual + b5*Ran + b6*Serv
$
```

A.2.8 – Instructions for survey participants for the main survey

School of Social Sciences
CITY UNIVERSITY LONDON

Instructions for Survey Participants

A very big thank you for agreeing to participate in this anonymous survey about the UK groceries market. Your honest contribution will be an important component towards the completion of PhD research.

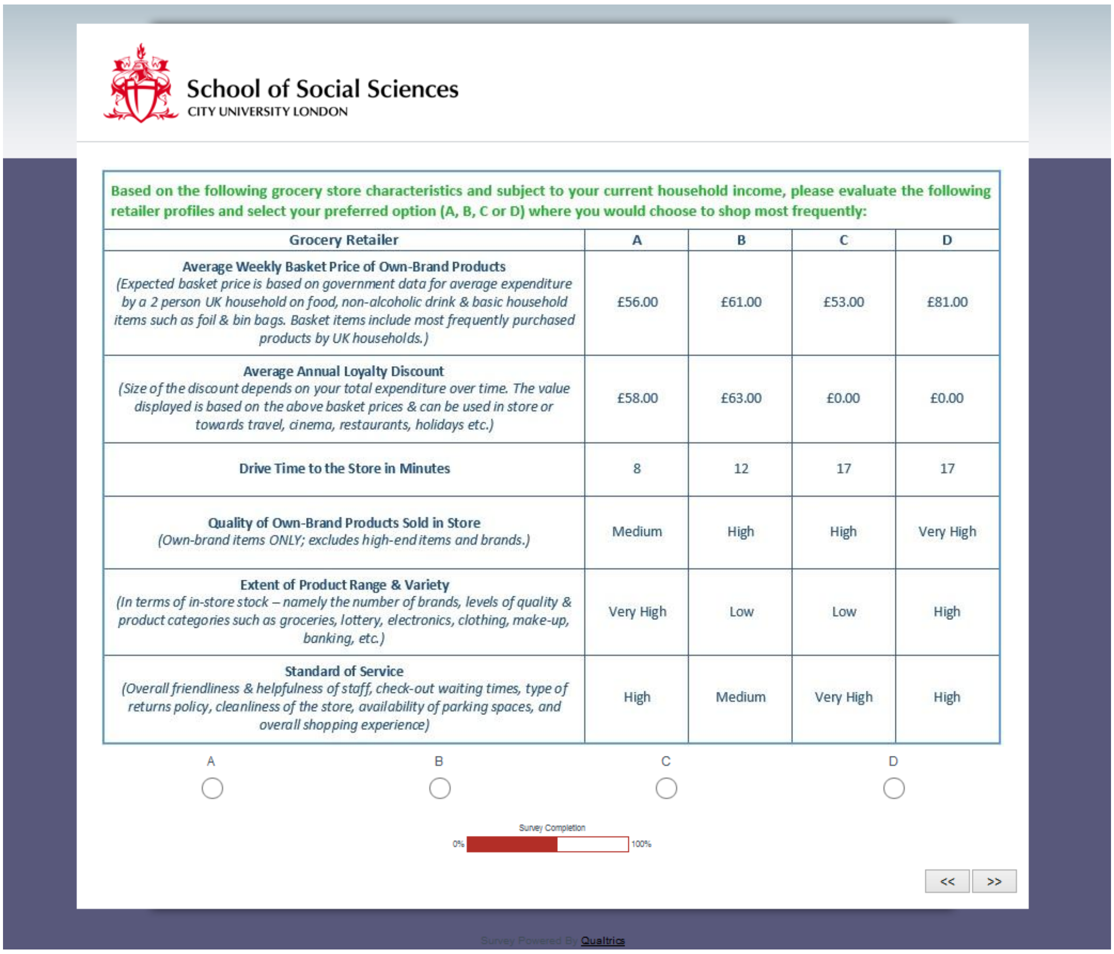
The following survey is made up of two sections and should take you around 10-15 minutes to complete.

Section 1 incorporates 10 multiple choice questions regarding your current household characteristics and shopping behaviour.

Section 2 consists of 11 identical choice situations asking you to evaluate and identify your preferred grocery retailer profile from 4 available options.These choice scenarios should be treated as hypothetical shopping trips.
Please note that only the values of the descriptive characteristics will change between the questions so please carefully assess these differences between the available options when making your choice.

You must evaluate each of the choice scenarios and descriptive characteristics as an individual grocery shopping trip.

A.2.9 – Screenshot of a sample question from the main survey



School of Social Sciences
CITY UNIVERSITY LONDON

Based on the following grocery store characteristics and subject to your current household income, please evaluate the following retailer profiles and select your preferred option (A, B, C or D) where you would choose to shop most frequently:

Grocery Retailer	A	B	C	D
Average Weekly Basket Price of Own-Brand Products *(Expected basket price is based on government data for average expenditure by a 2 person UK household on food, non-alcoholic drink & basic household items such as foil & bin bags. Basket items include most frequently purchased products by UK households.)*	£56.00	£61.00	£53.00	£81.00
Average Annual Loyalty Discount *(Size of the discount depends on your total expenditure over time. The value displayed is based on the above basket prices & can be used in store or towards travel, cinema, restaurants, holidays etc.)*	£58.00	£63.00	£0.00	£0.00
Drive Time to the Store in Minutes	8	12	17	17
Quality of Own-Brand Products Sold in Store *(Own-brand items ONLY; excludes high-end items and brands.)*	Medium	High	High	Very High
Extent of Product Range & Variety *(In terms of in-store stock – namely the number of brands, levels of quality & product categories such as groceries, lottery, electronics, clothing, make-up, banking, etc.)*	Very High	Low	Low	High
Standard of Service *(Overall friendliness & helpfulness of staff, check-out waiting times, type of returns policy, cleanliness of the store, availability of parking spaces, and overall shopping experience)*	High	Medium	Very High	High

A B C D

Survey Completion
0% 100%

<< >>

Survey Powered By Qualtrics

A.3.1 – List of Model Specifications

Model	Model specification
CL1	• Conditional logit • Contains only the main explanatory variables (no interaction terms) • Applies population weights • Individual-level clustered standard errors
CL1 (a)*	• Conditional logit • Contains only the main explanatory variables (no interaction terms) • Applies population weights
CL1 (b)*	• Conditional logit • Contains only the main explanatory variables (no interaction terms) • Individual-level clustered standard errors
CL2	• Conditional logit • Contains the main explanatory variables and all candidate interaction terms • Applies population weights • Individual-level clustered standard errors
CL3	• Conditional logit • Contains the main explanatory variables and statistically significant interaction terms • Applies population weights • Individual-level clustered standard errors
ML1	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • All explanatory variables are assumed to be random • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML1 (a)*	• Mixed logit interim model to identify which of the explanatory variables has significant standard deviations to verify if preferences vary in the population for that particular attribute • Variables assumed to be random and normally distributed are Price, Discount, Time, Very High Quality, High and Very High Range, High and Very High Service • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML1 (b)*	• Interim mixed logit to identify which of the explanatory variables has significant standard deviations to verify if preferences vary in the population for that particular attribute • Variables assumed to be random and normally distributed are Price, Discount, Time, Very High Quality, Very High Range, High and Very High Service • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML2	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights

	• Individual-level clustered standard errors • 500 Halton draws
ML2 (a)*	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • 500 Halton draws
ML2 (b)*	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Individual-level clustered standard errors • 500 Halton draws
ML2 (c)*	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • Individual-level clustered standard errors • 50 Halton draws
ML3	• Mixed logit • Contains the main explanatory variables and all candidate interaction terms • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML3 (a)*	• Mixed logit • Contains the main explanatory variables and all candidate interaction terms • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • Individual-level clustered standard errors • 50 Halton draws
ML4	• Mixed logit • Contains the main explanatory variables and statistically significant interaction terms • Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML4 (a)*	• Mixed logit • Contains the main explanatory variables and statistically significant interaction terms

	• Variables assumed to be random and normally distributed are restricted to Price, Discount, Time, Very High Quality, Very High Range and High Service • Applies population weights • Individual-level clustered standard errors • 50 Halton draws
ML5	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Discount, Very High Quality, Very High Range and High Service • Variables assumed to be random and log-normally distributed are restricted to Price and Time • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML6	• Mixed logit • Contains the main explanatory variables and significant interaction terms • Variables assumed to be random and normally distributed are restricted to Discount, Very High Quality, Very High Range and High Service • Variables assumed to be random and log-normally distributed are restricted to Price and Time • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML6 (a)*	• Mixed logit • Contains the main explanatory variables and all candidate interaction terms • Variables assumed to be random and normally distributed are restricted to Discount, Very High Quality, Very High Range and High Service • Variables assumed to be random and log-normally distributed are restricted to Price and Time • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML7 (a)	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Discount, Time, Very High Quality, Very High Range and High Service • Price is assumed to have a fixed coefficient for WTP estimates • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML7 (b)*	• Interim mixed logit to find significant interaction terms • Contains the main explanatory variables and interaction terms that were statistically significant in model ML4 • Variables assumed to be random and normally distributed are restricted to Discount, Time, Very High Quality, Very High Range and High Service • Price is assumed to have a fixed coefficient for WTP estimates • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML7 (c)	• Mixed logit

	• Contains the main explanatory variables and only statistically significant interaction terms • Variables assumed to be random and normally distributed are restricted to Discount, Time, Very High Quality, Very High Range and High Service • Price is assumed to have a fixed coefficient for WTP estimates • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML8 (a)	• Mixed logit • Contains only the main explanatory variables (no interaction terms) • Variables assumed to be random and normally distributed are restricted to Discount, Very High Quality, Very High Range and High Service • Variables assumed to be random and log-normally distributed are restricted to Time • Price is assumed to have a fixed coefficient for WTP estimates • Applies population weights • Individual-level clustered standard errors • 500 Halton draws
ML8 (b)	• Contains the main explanatory variables and significant interaction terms from ML 6 (b) Variables assumed to be random and normally distributed are restricted to Discount, Very High Quality, Very High Range and High Service • Price is assumed to have a fixed coefficient for WTP estimates • Variables assumed to be random and log-normally distributed are restricted to Time • Applies population weights • Individual-level clustered standard errors • 500 Halton draws

Notes: * indicates that the table of results for this model specification is presented in the appendix. This table does not include the pilot survey model specifications which are outlined in Chapter 2.

A.3.2 - Description of interaction terms tested in the conditional and mixed logit model specifications

Interaction	Interaction description
Female*Price	The effect of gender on preferences for grocery basket price.
Female*Discount	The effect of gender on preferences for a loyalty scheme discount.
Female*Time	The effect of gender on preferences for grocery store proximity.
Female*VH Quality	The effect of gender on preferences for very high quality of products at the grocery store.
Female*VH Range	The effect of gender on preferences for very high range of products at the grocery store.
Female*High Service	The effect of gender on preferences for high levels of service at the grocery store
Large Household*Price	The effect of living in a large household of four or more individuals, on preferences for grocery basket price.
Large Household*Discount	The effect of living in a large household of four or more individuals, on preferences for a loyalty scheme discount.
Large Household*VH Range	The effect of living in a large household of four or more individuals, on preferences for very high range of products at the grocery store.
Unemployed*Price	The effect of being unemployed on preferences for grocery basket price.
Student*Price	The effect of being a student on preferences for grocery basket price.
Unemployed*Discount	The effect of being unemployed on preferences for a loyalty scheme discount.
Student*Discount	The effect of being a student on preferences for a loyalty scheme discount.
Unemployed*Time	The effect of being unemployed on preferences for grocery store proximity.
Student*Time	The effect of being a student on preferences for grocery store proximity.
No Car*Time	The effect of not driving a car to go grocery shopping on grocery store proximity.
No Car*VH Range	The effect of not driving a car to go grocery shopping on very high product range.
18 – 44 Age Group*Price	The effect of being in the 18-44 age group on preferences for the grocery basket price.
18 – 44 Age Group*Discount	The effect of being in the 18-44 age group on preferences for a loyalty scheme discount.

< £22,000 HI*Price	The effect of being in the lower household income group earning less than £22,000 annually (after tax) on preferences for grocery basket price.
> £45,000 HI*Price	The effect of being in a higher household income group earning more than **£**45,000 annually (after tax) on preference for grocery basket price.
> £45,000 HI*Discount	The effect of being in a higher household income group earning more than **£**45,000 annually (after tax) on preference for the loyalty scheme discount.
> £45,000 HI*VH Quality	The effect of being in a higher household income group earning more than **£**45,000 annually (after tax) on preference for very high quality of products in store.
> £45,000 HI*High Service	The effect of being in a higher household income group earning more than **£**45,000 annually (after tax) on preference for very high levels of service in store.
> £45,000 HI*Time	The effect of being in a higher household income group earning more than **£**45,000 annually (after tax) on preference for grocery store proximity.
< £22,000 HI*Discount	The effect of being in a lower household income group earning less than £22,000 annually (after tax) on preference for the loyalty scheme discount.
Frequent Online Shop*Price	The effect of regularly purchasing groceries online (i.e. at least once a month) on preferences for the grocery basket price.
Frequent Online Shop*Discount	The effect of regularly purchasing groceries online (i.e. at least once a month) on preferences for the loyalty scheme discount.
Frequent Online Shop*Time	The effect of regularly purchasing groceries online (i.e. at least once a month) on preferences for grocery store proximity.
Infrequent Online Shop*Price	The effect of rarely purchasing groceries online (i.e. few times a year/ never) on preferences for the grocery basket price.
Infrequent Online Shop*Discount	The effect of rarely purchasing groceries online (i.e. few times a year/ never) on preferences for the loyalty scheme discount.
Infrequent Online Shop*Time	The effect of rarely purchasing groceries online (i.e. few times a year/ never) on preferences for grocery store proximity.
Infrequent Online Shop*VH Range	The effect of rarely purchasing groceries online (i.e. few times a year/ never) on preferences for very high product range in store.
No Loyalty Cards*Price	The effect of not participating in any loyalty schemes on preferences for the grocery basket price.
No Loyalty Cards*Discount	The effect of not participating in any loyalty schemes on preference for the loyalty scheme discount.
No Loyalty Cards*Time	The effect of not participating in any loyalty schemes on preferences for grocery store proximity.

1-2 Loyalty Cards*Time	The effect of participating in 1-2 loyalty schemes on preferences for grocery store proximity.
1-2 Loyalty Cards*Discount	The effect of participating in 1-2 loyalty schemes on preferences for the loyalty scheme discount.
1-2 Loyalty Cards*Price	The effect of participating in 1-2 loyalty schemes on preferences for the grocery basket price.

A.3.3 – Conditional Logit Models CL1 (a) & CL1 (b)

	CL1 (a)			CL1 (b)		
Variable	**Coef.**	**|*z*-stat|**	**O.R.**	**Coef.**	**|*z*-stat|**	**O.R.**
Price	-0.054***			-0.054***		
	(0.002)	-23.94	0.947	(0.003)	-15.95	0.947
Discount	0.009***			0.009***		
	(0.001)	16.22	1.009	(0.001)	9.66	1.009
Time	-0.061***			-0.075***		
	(0.005)	-12.84	0.941	(0.006)	-12.58	0.928
Medium Quality	0.558***			0.725***		
	(0.061)	9.09	1.746	(0.071)	10.21	2.065
High Quality	0.555***			0.764***		
	(0.065)	8.59	1.741	(0.082)	9.31	2.146
Very High Quality	0.830***			1.035***		
	(0.075)	11.05	2.292	(0.089)	11.62	2.815
Medium Range	0.707***			0.571***		
	(0.065)	10.82	2.027	(0.062)	9.24	1.769
High Range	0.924***			0.960***		
	(0.070)	13.18	2.519	(0.072)	13.27	2.613
Very High Range	1.047***			1.006***		
	(0.061)	17.12	2.850	(0.068)	14.7	2.735
Medium Service	0.845***			0.782***		
	(0.059)	14.35	2.327	(0.062)	12.6	2.186
High Service	0.988***			0.945***		
	(0.058)	17.1	2.687	(0.072)	13.11	2.572
Very High Service	1.152***			1.110***		
	(0.073)	15.79	3.165	(0.081)	13.68	3.034
Log-likelihood	-5204.184			-5109.258		
Nr. of Resp.	427			427		
Nr. of Obs.	18832			18788		

Notes: Standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

A.3.4 – Mixed Logit Models ML1 (a) & ML1 (b)

	ML1 (a)			ML1 (b)		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev**
Price	-0.094*** (0.010)	9.35	0.067*** (0.008)	-0.091*** (0.010)	8.97	0.067*** (0.009)
Discount	0.012*** (0.002)	5.24	0.024*** (0.003)	0.012*** (0.002)	5.11	-0.025*** (0.003)
Time	-0.107*** (0.015)	7.3	0.110*** (0.013)	-0.104*** (0.015)	7.09	0.111*** (0.015)
Medium Quality	0.473*** (0.156)	3.02	0.392** (0.178)	0.478*** (0.161)	2.97	0.286 (0.544)
High Quality	0.563*** (0.193)	2.92	0.284 (0.283)	0.574*** (0.192)	2.99	-
Very High Quality	0.875** (0.188)	4.64	0.688*** (0.206)	0.836*** (0.197)	4.25	0.659*** (0.235)
Medium Range	0.741*** (0.136)	5.45	-	0.743*** (0.137)	5.43	-
High Range	0.982*** (0.165)	5.94	-	0.983*** (0.164)	5.98	-
Very High Range	1.196*** (0.158)	7.58	0.863*** (0.132)	1.209*** (0.156)	7.77	0.893*** (0.147)
Medium Service	1.141*** (0.161)	7.08	-	1.122*** (0.154)	7.27	-
High Service	1.320*** (0.170)	7.77	0.679*** (0.171)	1.330*** (0.163)	8.15	0.757*** (0.148)
Very High Service	1.433*** (0.211)	6.79	-	1.416*** (0.205)	6.9	-
Log-likelihood	-4640.753			-4646.057		
Nr. of Resp.	427			427		
Nr. of Obs.	18832			18832		

Notes: Robust standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

A.3.5 – Mixed Logit Models ML2 (a), ML2 (b) & ML2 (c)

	ML2 (a)			ML2 (b)			ML2 (c)		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev**	**Coef.**	**|*z*-stat|**	**St. Dev**
Price	-0.090***		0.067***	-0.089***	16.21	-0.065***	-0.080***	7.18	0.060**
	(0.005)	16.99	(0.005)	(0.005)		(0.006)	(0.011)		(0.024)
Discount	0.011***		0.024***	0.012***	8.96	-0.023***	0.011***	4.49	0.024***
	(0.001)	7.98	(0.001)	(0.001)		(0.002)	(0.003)		(0.003)
Time	-0.102***		0.106***	-0.126***	12.99	0.131***	-0.092***	7.09	0.100***
	(0.008)	12.63	(0.008)	(0.010)		(0.010)	(0.013)		(0.015)
Medium	0.475***			0.701***	7.72		0.488***	3.20	-
Quality	(0.073)	6.49	-	(0.091)		-	(0.153)		
High Quality	0.573***			0.843***	7.68		0.580***	3.15	-
	(0.079)	7.28	-	(0.110)		-	(0.184)		
Very High	0.840***		0.696***	1.087***	9.98	0.665***	0.809***	4.39	0.527**
Quality	(0.094)	8.94	(0.106)	(0.109)		(0.106)	(0.184)		(0.256)
Medium	0.740***			0.513***	7.05		0.711***	5.53	-
Range	(0.076)	9.76	-	(0.073)		-	(0.129)		
High Range	0.984***			1.018***	11.51		0.955***	6.08	-
	(0.080)	12.3	-	(0.088)		-	(0.157)		
Very High	1.201***		0.882***	1.110***	12.54	0.900***	1.151***	7.75	0.709***
Range	(0.087)	13.87	(0.083)	(0.088)		(0.087)	(0.149)		(0.210)
Medium	1.127***			1.005***	12.00		1.086***	7.35	-
Service	(0.072)	15.57	-	(0.084)		-	(0.148)		
High Service	1.340***		0.721***	1.211***	13.47	-0.692***	1.229***	7.85	0.646***
	(0.079)	16.90	(0.085)	(0.090)		(0.101)	(0.157)		(0.160)
Very High	1.403***			1.321***	12.51		1.367***	6.83	-
Service	(0.087)	16.15	-	(0.106)		-	(0.200)		
Log-likelihood	-4647.902			-4600.321			-4674.851		
Nr. of Resp.	427			427			427		
Nr. of Obs.	18,832			18,788			18,832		

Notes: Standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

A.3.6 – Mixed logit models ML3 (a) & ML4 (a)

	ML3 (a)			ML4 (a)		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev.**
Price	-0.076*** (0.019)	4.1	0.062*** (0.009)	-0.076 (0.016)	4.86	0.051*** (0.010)
Discount	0.025*** (0.007)	3.78	0.024*** (0.003)	0.011 (0.003)	4.15	0.024*** (0.003)
Time	-0.158*** (0.032)	4.95	0.091*** (0.016)	-0.121 (0.014)	8.45	0.091*** (0.013)
Medium Quality	0.488*** (0.156)	3.13	-	0.499 (0.154)	3.23	-
High Quality	0.588*** (0.195)	3.02	-	0.602 (0.190)	3.17	-
Very High Quality	1.125*** (0.244)	4.61	0.475** (0.213)	1.137 (0.238)	4.78	0.443** (0.192)
Medium Range	0.725*** (0.129)	5.62	-	0.708 (0.126)	5.60	-
High Range	0.947*** (0.156)	6.06	-	0.948 (0.156)	6.09	-
Very High Range	0.990*** (0.261)	3.79	0.739*** (0.161)	1.000 (0.162)	6.17	0.693*** (0.170)
Medium Service	1.122*** (0.158)	7.12	-	1.088 (0.153)	7.11	-
High Service	1.056*** (0.180)	5.86	0.679*** (0.198)	1.156 (0.165)	7.00	0.667*** (0.148)
Very High Service	1.408*** (0.208)	6.76	-	1.376 (0.205)	6.71	-
Female*Price	-0.019* (0.011)	1.76	-	-	-	-
Female*Discount	0.003 (0.005)	0.54	-	-	-	-
Female*Time	-0.014 (0.025)	0.57	-	-	-	-
Female*VH Quality	-0.540** (0.233)	2.32	-	-0.565** (0.232)	2.43	
Female*VH Range	0.027 (0.221)	0.12	-	-	-	-
Female*High Service	0.144 (0.190)	0.75	-	-	-	-
Large Household*Price	0.027** (0.011)	2.56	-	-	-	-
Large Household*Discount	0.006 (0.005)	1.16	-	0.006 (0.006)	1.12	
Large Household*VH Range	-0.226 (0.229)	0.99	-	-	-	-
Unemployed*Price	0.006 (0.026)	0.23	-	-	-	-
Student*Price	0.025* (0.015)	1.73	-	-	-	-
Unemployed*Discount	-0.003 (0.006)	0.48	-	-	-	-
Student*Discount	-0.005*** (0.005)	0.9	-	-	-	-
Unemployed*Time	0.032 (0.041)	0.78	-	-	-	-
Student*Time	-0.037	1.15	-	-	-	-

	(0.032)					
No Car*Time	-0.022 (0.022)	0.97	-	-	-	-
No Car*VH Range	0.340 (0.232)	1.46	-	0.375 (0.241)	1.56	-
18 – 44 Age Group*Price	-0.027** (0.012)	2.2	-	-		-
18 – 44 Age Group*Discount	-0.004 (0.005)	0.84	-	-	-	-
< £22,000 HI*Price	-0.019 (0.012)	1.59	-	-0.033* (0.019)	1.72	-
> £45,000 HI*Price	0.023 (0.017)	1.36	-	-	-	-
> £45,000 HI*Discount	-0.006 (0.007)	0.86	-	-	-	-
> £45,000 HI*VH Quality	0.317 (0.299)	1.06	-	-	-	-
> £45,000 HI*High Service	0.562** (0.233)	2.41	-	0.521** (0.261)	2.00	-
< £22,000 HI*Discount	-0.007* (0.004)	1.68	-	-	-	-
> £45,000 HI*Time	0.043 (0.031)	1.41	-	-	-	-
Frequent Online Shop*Price	0.059*** (0.011)	5.17	-	0.046** (0.021)	2.2	-
Frequent Online Shop*Discount	0.000 (0.004)	0.08	-	-	-	-
Frequent Online Shop*Time	0.101*** (0.023)	4.46	-	0.082*** (0.023)	3.59	-
Infrequent Online Shop*Price	-0.011 (0.011)	0.98	-	-	-	-
Infrequent Online Shop*Discount	-0.005 (0.004)	1.38	-	-	-	-
Infrequent Online Shop*Time	0.024 (0.023)	1.03	-	-	-	-
Infrequent Online Shop*VH Range	0.088 (0.244)	0.36	-	-	-	-
No Loyalty Cards*Price	0.020 (0.013)	1.56	-	-	-	-
No Loyalty Cards*Discount	-0.017*** (0.007)	2.58	-	-0.014*** (0.004)	3.19	-
No Loyalty Cards*Time	0.060** (0.030)	1.98	-	-	-	-
1-2 Loyalty Cards*Time	0.034 (0.023)	1.43	-	-	-	-
1-2 Loyalty Cards*Discount	-0.004 (0.005)	0.72	-	-	-	-
1-2 Loyalty Cards*Price	-0.003 (0.011)	0.24	-	-	-	-
Log-likelihood	-4537.429			-4592.741		
Nr. of Resp.	427			427		
Nr. of Obs.	18832			18832		

Notes: Robust standard errors presented in parentheses. ***, **, *, next to coefficients represents statistical significance at the 1%, 5%, and 10% respectively.

A.3.7 – Mixed logit models ML6 (a) & ML7 (b)

	ML6 (a)			ML7 (b)		
Variable	**Coef.**	**|*z*-stat|**	**St. Dev.**	**Coef.**	**|*z*-stat|**	**St. Dev**
Price	-0.112*** (0.021)	5.31	0.125*** (0.032)	-0.058*** (0.010)	-5.81	0.024*** (0.003)
Discount	0.025*** (0.008)	3.29	0.023*** (0.003)	0.011*** (0.003)	4.18	0.102*** (0.014)
Time	-0.155*** (0.035)	4.37	0.174*** (0.064)	-0.122*** (0.015)	-8.35	-
Medium Quality	0.457*** (0.162)	2.82	-	0.534*** (0.145)	3.68	-
High Quality	0.551*** (0.197)	2.8	-	0.621*** (0.181)	3.43	-
Very High Quality	1.116*** (0.254)	4.39	0.578*** (0.187)	1.224*** (0.242)	5.06	0.611*** (0.187)
Medium Range	0.700*** (0.133)	5.24	-	0.752*** (0.131)	5.74	-
High Range	0.937*** (0.161)	5.8	-	0.984*** (0.155)	6.35	-
Very High Range	1.020*** (0.277)	3.68	0.879*** (0.130)	1.134*** (0.170)	6.69	0.976*** (0.122)
Medium Service	1.129*** (0.157)	7.17	-	1.077*** (0.149)	7.22	-
High Service	1.091*** (0.180)	6.04	0.683*** (0.166)	1.285*** (0.167)	7.7	0.703*** (0.129)
Very High Service	1.410*** (0.210)	6.71	-	1.421*** (0.199)	7.14	-
Female*Price	-0.006 (0.013)	0.42		-	-	-
Female*Discount	0.002 (0.005)	0.33		-	-	-
Female*Time	-0.049** (0.026)	1.87		-	-	-
Female*VH Quality	-0.559** (0.240)	2.32		-0.666*** (0.241)	2.76	-
Female*VH Range	0.028 (0.237)	0.12		-	-	-
Female*High Service	0.183 (0.205)	0.89		-	-	-
Large Household*Price	0.026** (0.011)	2.41		-	-	-
Large Household*Discount	0.007 (0.005)	1.37		0.005 (0.005)	1.07	-
Large Household*VH Range	-0.160 (0.236)	0.68		-	-	-
Unemployed*Price	-0.035** (0.016)	2.15		-	-	-
Student*Price	0.008 (0.013)	0.59		-	-	-
Unemployed*Discount	0.002 (0.007)	0.31		-	-	-
Student*Discount	-0.001 (0.005)	0.22		-	-	-
Unemployed*Time	0.018 (0.032)	0.58		-	-	-
Student*Time	-0.042 (0.029)	1.45		-	-	-
No Car*Time	-0.007 (0.021)	0.35		-	-	-

No Car*VH Range	0.411* (0.236)	1.74	0.369 (0.253)	1.46	-
18 – 44 Age Group*Price	-0.019 (0.013)	1.49	-	-	-
18 – 44 Age Group*Discount	-0.001 (0.005)	0.33	-	-	-
< £22,000 HI*Price	-0.021* (0.012)	1.8	-0.035*** (0.011)	3.23	-
> £45,000 HI*Price	0.034* (0.019)	1.79	-	-	-
> £45,000 HI*Discount	-0.008 (0.007)	1.23	-	-	-
> £45,000 HI*VH Quality	0.383 (0.294)	1.3	-	-	-
> £45,000 HI*High Service	0.554** (0.233)	2.38	0.365* (0.210)	1.74	
< £22,000 HI*Discount	-0.007 (0.005)	1.62	-	-	-
> £45,000 HI*Time	0.023 (0.037)	0.62	-	-	-
Frequent Online Shop*Price	0.045*** (0.012)	3.73	0.036*** (0.011)	3.15	
Frequent Online Shop*Discount	0.001 (0.005)	0.27	-	-	-
Frequent Online Shop*Time	0.087*** (0.021)	4.12	0.084*** (0.022)	3.89	
Infrequent Online Shop*Price	-0.005 (0.013)	0.37	-	-	-
Infrequent Online Shop*Discount	-0.003 (0.004)	0.66	-	-	-
Infrequent Online Shop*Time	0.026 (0.022)	1.15	-	-	-
Infrequent Online Shop*VH Range	0.021 (0.250)	0.09	-	-	-
No Loyalty Cards*Price	0.010 (0.013)	0.75	-	-	-
No Loyalty Cards*Discount	-0.022*** (0.007)	3.15	-0.018*** (0.005)	3.88	-
No Loyalty Cards*Time	0.040 (0.028)	1.44	-	-	-
1-2 Loyalty Cards*Time	0.028 (0.022)	1.25	-	-	-
1-2 Loyalty Cards*Discount	-0.008 (0.005)	1.51	-	-	-
1-2 Loyalty Cards*Price	0.009 (0.013)	0.72	-	-	-
Log-likelihood	-4511.693		-4667.179		
Nr. of Resp.	427		427		
Nr. of Obs.	18,832		18,832		

Notes: Robust standard errors presented in parentheses. ***, **, *, represent statistical significance at the 1%, 5%, and 10% respectively.

Bibliography

1 Achtnicht, M., Bühler, G., Hermeling, C., ‘The impact of fuel availability on demand for alternative-fuel vehicles’, *Transportation Research Part D: Transport and Environment*, Vol. 17, Issue 3, 2012, pp. 262-269.

2 Ai, C., Norton, E.C., ‘Interaction Terms in Logit and Probit Models’, *Economics Letters*, No. 80, 2003, pp. 123-129.

3 Amelio, A., Donath, D., ‘Market Definition in Recent EC Merger Investigations: The role of Empirical Analysis’, *Concurrences - Law & Economics,* No. 3, 2009.

4 Anderson, E.T., Kumar, N., ‘Price Competition with Repeat, Loyal Buyers’, *Quantitative Marketing and Economics*, No. 5, 2007, pp. 333-359.

5 Asplund, M., Eriksson R., Strand, N., ‘Price Discrimination in Oligopoly: Evidence from Regional Newspapers’, *The Journal of Industrial Economics,* No. 2, Vol. LVI, June 2008.

6 Avery, R.J., ‘Determinants of Search for Nondurable Goods: An Empirical Assessment of the Economics of Information Theory’, *The Journal of Consumer Affairs,* Vol. 30, No. 2, 1996, pp. 390-420.

7 Basso, L.J., Clements M.T., Ross, T.W., ‘Moral Hazard and Customer Loyalty Programs’, *American Economic Journal: Microeconomics*, Vol. 1, Nr. 1, 2009, pp. 101-123.

8 Bately, R., “On Ordinal Utility, Cardinal Utility and Random Utility”, *Theory and Decision*, Vol. 64, Issue 1, 2008, pp. 37-63.

9 Bech, M., Gyrd-Hansen, D., ‘Effects Coding in Discrete Choice Experiments’, *Health Economics*, No. 14, 2005, pp. 1079-1083.

10 Bech, M., Kjaer, T., Lauridsen, J., ‘Does the Number of Choice Sets Matter? Results from a Web Survey Applying a Discrete Choice Experiment’, *Health Economics*, No. 20, 2011, pp. 273-286.

11 Beggs, A., Klemperer, P., 'Multi-Period Competition with Switching Costs', *Econometrica,* Vol. 60, Nr. 3, 1992, pp. 651-666.

12 Bekker-Grob, E,W., Ryan, M., Gerard, K., ‘Discrete choice experiments in health economics: a review of the literature’, *Health Economics,* Volume 21, Issue 2, 2012, pp. 145-172.

13 Biglaiser, G., Crémer J., Dobos, G., ‘Heterogeneous switching costs’, International Journal of Industrial Organization, Vol. 47, 2016, pp. 62-87.

14 Blass, A.A., Lach, S., Manski, C.F., ‘Using Elicited Choice Probabilities to Estimate Random Utility Models: Preferences for Electricity Reliability’, *International Economic Review*, Vol. 51, No. 2, May 2010.

15 Bliemer, M.C.J., Rose, J.M., ‘Confidence intervals of willingness-to-pay for random coefficient logit models’, *Transportation Research Part B*, 58, 2013, pp.199-214.

16 Bliemer, M.C.J., Rose, J.M., ‘Constructing Efficient Stated Choice Experimental Designs’, *Transport Reviews*, Vol. 29, No. 5, 2009, pp. 587-617.

17 Bonaria Lai, M., Del Giudice, T., Pomarici, E., ‘Unoberserved Heterogeneity in the Wine Market : an Analysis of Sardinian Wine using Mixed Logit’, AAWE Working Paper, American Association of Wine Economists, No. 28, 2008.

18 Caminal, R., ‘The Design and Efficiency of Loyalty Rewards’, *Journal of Economics and Management Strategy*, Vol. 21, Nr. 2, 2012, pp. 339-371.

19 Caminal, R., Calici, A., ‘Are loyalty-rewarding pricing schemes anti-competitive?’ *International Journal of Industrial Organization*, Vol. 25, 2007, pp. 657-674.

20 Caminal, R., Matutes, C., 'Endogenous Switching Costs in a Duopoly Model', *International Journal of Industrial Organisation*, 8, 1990, pp. 353-373.

21 Campbell, Cora, ‘The Northern Ireland Retail Landscape’, Kantar Worldpanel, 27th June 2017, https://www.food.gov.uk/sites/default/files/coracampbellkantarworldpanel.pdf, (accessed 09/10/2017).

22 Carson, R., Pihlens, D., Louviere, J.J., ‘Design of Choice Experiments: A Discussion of Issues That Matter in Future Applied Research', *Journal of Choice Modelling*, 4(1), 2011, pp 1-8.

23 ChoiceMetrics, Forum Support, www.choice-metrics.com/forums

24 ChoiceMetrics, Ngene 1.1.1 User Manual & Reference Guide, Australia, www.choice-metrics.com, 2012.

25 Chung, J., Rao, V., ‘A General Choice Model for Bundles with Multiple-Category Products: Application to Market Segmentation and Optimal Pricing for Bundles’, *Journal of Marketing Research*, Vol. XL, 2003, pp. 115-130.

26 Colomé, R., Serra A., ‘Supermarket key attributes and location decisions: A comparative study between British and Spanish consumers’, *Economics Working Papers*, No. 469, Department of Economics and Business, Universitat Pompeu Fabra, 2010.

27 Competition Commission, Groceries Market Investigation, *Final Report and Annexes*, 2008.

28 Competition Commission, Review of Stated Preference and Willingness to Pay Methods, *Review Report*, 2010.

29 Competition & Markets Authority, Energy market investigation, *Final Report*, 2016.

30 Competition & Markets Authority, Retail banking market investigation, *Final Report*, 2016.

31 Dahlberg M., Eklof, M., Fredriksson, Jofre-Monseny, J., 'Estimating Preferences for Local Public Services Using Migration Data', *Urban Studies Journal*, 2012.

32 Davies, S., Thomas, D., 'Mobile payments groups set sights on winning over UK wallets', *The Financial Times*, June 2, 2014.

33 DeGraba, P., 'The loss leader is a turkey: Targeted discounts from multi-product competitors', *International Journal of Industrial Organization*, 24, 2006, pp. 613–628.

34 Doherty, E., Campbell, D.**,** Hynes, S. and van Rensburg, T.M, 'Examining Labelling Effects Within Discrete Choice Experiments: An Application to Recreational Site Choice', *Journal of Environmental Management*, Nr 125, 2013, pp. 94-104.

35 Dorotic M., Bijmolt T., Verhoef P., 'Loyalty Programmes: Current Knowledge and Research Directions', *International Journal of Management Reviews*, 14(3), September 2012, pp. 217-237.

36 Economides, N., 'Tying, Bundling and Loyalty Requirement Rebates', Working Paper 10-26, Stern School of Business, NYU, 2010.

37 Economist, The, 'Loyalty Rewards and Insurance: Every little Helps', *The Economist*, November 5, 2011.

38 Economist, The, 'Retailing: Spies in your Wallet', *The Economist*, November 5, 2011.

39 Elhauge, E., 'Tying, Bundled Discounts, And the Death of the Single Monopoly Profit Theory', *Harvard Law Review*, Vol. 123, No. 2, 2009 .

40 European Commission, *Guidance on the Commission's Enforcement priorities in Applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings*, 2009/C 45/02, 2009.

41 Faella, G., 'The Antitrust Assessment of Loyalty Discounts and Rebates', *Journal of Competition Law & Economics*, 4(2), 2008, pp. 375-410.

42 Farrell, J., Klemperer, P., 'Coordination and Lock-In: Competition with Switching Costs and Network Effects', *Handbook of Industrial Organizatio*, Chapter 31, Volume 3, 2007.

43 Federico, G., 'The Antitrust Treatment of Loyalty Discounts in Europe: Towards a more Economic Approach', *Journal of European Law & Practice*, Vol. 2, No. 3, 2011.

44 Fong, Y.F., Liu, Q., 'Loyalty Rewards Facilitate Collusion', *Journal of Economics and Management Strategy*, Vol. 20, Nr. 3, 2011, pp. 739-775.

45 Fudenburg, D., Tirole, J., 'Customer Poaching & Brand Switching', *The Rand Journal of Economics*, Vol. 31, No. 4, Winter, 2000, pp. 634-657.

46 Gans, J., King, S., ‘Paying for Loyalty: Product Bundling in Oligopoly’, *Journal of Industrial Organization*, Vol. LIV, Nr. 1, March 2006.

47 Greene, W., ‘Testing Hypotheses about Interaction Terms in Nonlinear Models’, *Economics Letters*, No. 107, 2010, pp. 291-296.

48 Greene, W.H., *Econometric Analysis*, 6th Edition, Pearson-Prentice Hall, 2008.

49 Greenlee, P., Reitman, D., Sibley D.S., ‘An Anti-Trust Analysis of Bundled Loyalty Discounts’, *International Journal of Industrial Organization*, Vol. 26, 2008, pp. 1132-1152.

50 Guardian, The, ‘Sainsbury’s Nectar points cut angers customers’, 10 April 2015, http://www.theguardian.com/money/2015/apr/10/sainsburys-cuts-to-nectar-points-anger-its-customers, (accessed 1 June 2015)

51 Hackbarth, A., Madlener, R., ‘Consumer preferences for alternative fuel vehicles: A discrete choice analysis’, *Transportation Research Part D*, Vol. 25, 2013, pp. 5-17.

52 Hartmann, W.R., Viard, V.B. ‘Do Frequency Reward Programs Create Switching Costs?’, *Quantitative Marketing & Economics*., Vol. 6, 2008, pp. 109-137.

53 Hensher, D., Greene, W.H., ‘The Mixed Logit: The State of Practice’, *Transportation*, Vol. 30, Nr. 2, 2003, pp. 133-176.

54 Hess, S., Hensher, D., Daly, A., ‘Not Bored Yet - Revisiting Respondent Fatigue in Stated Choice Experiments’, *Transportation Research Part A*, No. 46, 2012, pp. 626-644.

55 Hess, S., Polak, J.W., ‘Mixed Logit Modelling of Airport Choice in Multi-Airport Regions’, *Journal of Air Transport Management*, Vol. 11, 2005, pp. 59-68.

56 Hole, A.R., ‘Fitting Mixed Logit Models by Using Maximum Simulated Likelihood’, *The Stata Journal*, 7, Nr. 3, 2007, pp. 388-401.

57 Hole, A.R., ‘Mixed Logit Modelling in Stata – an Overview’, Presentation for the UK Stata Users Group meeting, September 2013.

58 Hole, A.R., ‘Modelling Heterogeneity in Patients’ Preferences for the Attributes of a General Practitioner Appointment’, *Journal of Health Economics*, Vol. 27, 2008, pp. 1078-1094.

59 Hole, A.R., 'A comparison of approaches to estimating confidence intervals for willingness to pay measures', Health Economics, Vol. 16, Issue 8, 2007 pp. 827-840.

60 Huber, J., Zwerina, K., ‘The Importance of Utility Balance in Efficient Choice Designs’, *American Marketing Association*, Vol. 33, Nr. 3, 1996, pp. 307-317.

61 Hviid, M., ‘Supermarket Price Guarantees: and then there were Three’, UEA’s Centre for Competition Policy, 2011, http://competitionpolicy.wordpress.com/2011/10/14/supermarket-price-guarantees-%E2%80%93-and-then-there-were-three/, (accessed 1 December 2014)

62 IGD, ‘Shopper loyalty in 2015’, ShopperVista Report, April 2015, http://shoppervista.igd.com/Hub.aspx?id=32&tid=4&rptid=317, (accessed 1st October 2015)

63 IGD, ‘Top Up Shopping’, ShopperVista Report, June 2015, http://shoppervista.igd.com/Hub.aspx?id=32&tid=4&rptid=319, (accessed 1st October 2015)

64 IGD, ‘What impact do loyalty schemes have on store choice?’ 15th July 2013, http://www.igd.com/Research/Shopper-Insight/shopper-outlook/15151/What-impact-do-loyalty-schemes-have-on-store-choice/, (accessed 1st October 2015)

65 Keane, M., Wasi, N., ‘Comparing Alternative Models of Heterogeneity in Consumer Choice Behaviour, *Journal of Applied Econometrics*, Vol. 28, 2013, pp. 1018-1045.

66 Klemperer, P., 'Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade', *The Review of Economic Studies*, Vol. 62, No. 4, October, 1995, pp. 515-539.

67 Klemperer, P., 'Markets with consumer switching costs', *The Quarterly Journal of Economics,* Vol. 102, No. 2, May, 1987, pp. 375-394.

68 Lal, R., Bell, D.E., ‘The Impact of Frequent Shopper Programs in Grocery Retailing’ *Quantitative Marketing and Economics*, Vol. 1, 2003, pp. 179-202.

69 Lancaster, K.J., ‘A New Approach to Consumer Theory’, *Journal of Political Economy*, Vol. 74, No. 2, 1966, pp. 132-157.

70 Le Cadre, H., Bouhtou, M., Tuffin, B., ‘Consumers’ Preference Modelling to Price Bundle Offers in Telecommunication Industry: Study for a Single Provider’, *Netnomics*, Vol. 10, Nr. 2, 2009, pp.171-208.

71 Lederman, M., ‘Do Enhancements to Loyalty Programs Affect Demand? The Impact of International Frequent Flyer Partnerships on Domestic Airline Demand’, *The RAND Journal of Economics*, Vol. 37, No. 4, 2007, pp. 1134-1158.

72 Lianos, I., Genakos, C., ‘Econometric Evidence in EU Competition Law: an Empirical and Theoretical Analysis’, *Centre for Law, Economics and Society,* CLES Working Paper Series Nr. 6, UCL, 2012.

73 Liu, Y., Yang, R., ‘Competing Loyalty Programs: Impact of Market Saturation, Market Share, and Category Expandability’, *Journal of Marketing*, Vol. 73, 2009, pp. 93-108.

74 Louviere J.J., Hensher, D.A., Swait, J.D., *Stated Choice Methods: Analysis and Application,* Cambridge University Press, 2000.

75 Marti, J., ‘Assessing preferences for improved smoking cessation medications: a discrete choice experiment’, European Journal of Health Economics, Volume 13, Issue 5, 2012, pp 533-548.

76 Matutes, C., Regibeau, P., ‘Compatibility and Bundling of Complementary Goods in a Duopoly’, *The Journal of Industrial Economics*, Vol. 40, No. 1, 1992, pp. 37-54.

77 McCaughey, N.C., Behrens, C., ‘Paying for Status? The effect of Frequent Flier Program Member Status on Air Fare Choice’, *Discussion Paper 04/11*, Monash University, 2011.

78 McFadden, D., ‘Conditional Logit Analysis of Qualitative Choice Behaviour’, in P. Zarembka (ed.), *Frontiers of Econometrics*, Academic Press: New York, 1974, pp. 105-142.

79 Morell, A., Glockner, A., Towfigh, E., ‘Sticky Rebates: Target Rebates Induce Non-Rational Loyalty in Consumers’, *Max Planck Institute for Research on Collective Goods*, Nr. 23, Bonn, 2009.

80 Nichols, A., Schaffer, M., ‘Clustered Errors in Stata’, *Stata meeting presentation*, 10th September 2007, https://www.stata.com/meeting/13uk/nichols_crse.pdf, (accessed 24/10/2017).

81 Norton, E.C., Karaca-Mandic, P., Dowd, B., ‘Interaction Terms in Nonlinear Models’, *Health Services Research*, 47:1, 2012.

82 Pauler, G., Dick, A., ‘Maximising Profit of Food Retailing Chain by Targeting and Promoting Valuable Customers using Loyalty Card and Scanner Data’, *European Journal of Operational Research*, Nr. 174, 2006, pp. 1260-1280.

83 Quan, W., Rose, J.M., Collins, A.T., Bliemer, M.C.J., ‘A Comparison of Algorithms for Generating Efficient Choice Experiments’, *Institute of Transport and Logistics Studies*, Working Paper, ITLS-WP-11-19, 2011.

84 Regier, D.A., Ryan, M., Phimister, E., Marra C.A., ‘Bayesian and Classical Estimation of Mixed Logit: An Application to Genetic Testing’, *Journal of Health Economics*, Nr 28, 2009, pp.498-610.

85 Rhodes, A., 'Re-examining the effects of switching costs', *Economic Theory*, Vol. 57, Issue 1, 2014, pp. 161-194.

86 Rose, J.M., Scarpa, R., ‘Design Efficiency for Non-Market Valuation with Choice Modelling: How to Measure it, what to report and why’, *The Australian Journal of Agricultural and Resource Economics*, 52, 2008, pp. 253-282.

87 Rose, J.M., Scarpa, R., Bliemer M.C.J., ‘Incorporating Model Uncertainty into the Generation of Efficient Stated Choice Experiments: A Model Averaging Approach’, *Institute of Transport and Logistics Studies*, Working Paper, ITLS-WP-09-08, 2009.

88 Rowley, J., ‘Reconceptualising the strategic role of loyalty schemes’, *Journal of Consumer Marketing*, Vol. 24, Issue 7, 2007.

89 Sándor, Z., Wedel, M., ‘Profile Construction in Experimental Designs for Mixed Logit’, *Marketing Science*, Vol. 21, No. 4, 2002, pp. 455-475.

90 Shi, M., 'A theoretical analysis of endogenous and exogenous switching costs', *Quantitative Marketing & Economics*, 11, 2013, pp. 205–230.

91 Shin, J., Kim, C.S., Lee, J., ‘Model for Studying Commodity Bundling with a Focus on Consumer Preference’, *TEMEP Discussion Paper*, No. 34, 2009.

92 STATA FAQ, ‘How do I interpret odds ratios in logistic regression?’, UCLA: Statistical Consulting Group, http://www.ats.ucla.edu/stat/mult_pkg/faq/general/odds_ratio.htm, (accessed 8 December 2014)

93 StataCorp., *Stata 13 Base Reference Manual*, College Station, TX: Stata Press, 2013.

94 StataCorp., *Stata Statistical Software: Release 11*, College Station, TX: StataCorp. LP, 2009.

95 Stucke, M., Allen Grunes, A., *Big Data and Competition Policy*, Oxford University Press, paragraph 4.26, 2016.

96 Stopher, P., *Collecting, Managing, and Assessing Data Using Sample Surveys*, Cambridge University Press, 2012.

97 The Guardian, ‘Tesco to accept Sainsbury's Brand Match money-off vouchers’, 11 April 2016, http://www.theguardian.com/money/2016/apr/11/tesco-to-accept-sainsburys-brand-match-money-off-vouchers, (accessed 1 May 2016)

98 Tirole, J., *The Theory of Industrial Organization*, MIT, 1988.

99 TNS, 'Report: Consumer Attitudes Towards Food Standards', *UK Food Standards Agency*, February 2007.

100 Train, K., *Discrete Choice Methods with Simulation*, Second Edition, Cambridge University Press, 2009.

101 Turner, J.J., Wilson, K., ‘Grocery loyalty: Tesco Clubcard and its impact on loyalty’, *British Food Journal*, Vol. 108 Issue: 11, 2006, pp.958-964.

102 Villas-Boas, J.M., ‘Dynamic Competition with Experience Goods’, *Journal of Economics and Management Strategy*, Vol. 15, Nr. 1, 2006, pp. 37-66.

103 Volpe, R.J., ‘Promotional Competition Between Supermarket Chains’ *Review of Industrial Organization*, No. 42, 2013, pp. 45-61.

104 Wang, Z., ‘Supermarket and Gasoline: An Empirical Study of Bundled Discount’, *Working Paper*, 2010.

105 Waterson, M., 'Structural versus Quasi-Experimental Approaches to Industrial Organization', *International Journal of the Economics of Business*, Vol. 21, No. 1, 2014, pp. 43–47.

106 Waterson, M., 'The role of consumers in competition and competition policy', *International Journal of Industrial Organization*, 21, 2003, pp. 129–150.

107 Whish, R., *Competition Law*, Oxford University Press, Sixth Edition, 2009.

108 YouGov UK, ‘British shoppers in love with loyalty cards’, *YouGov Report*, November 7 2013, https://yougov.co.uk/news/2013/11/07/british-shoppers-love-loyalty-cards/, (accessed 1st March 2016)

109 Zenger, H., ‘Loyalty Rebates and the Competitive Process’, *Journal of Competition Law & Economics*, Vol. 8 (4), 2012, pp. 717-768.

Data Sources

Kantar Worldpanel, 'Accelerating the Growth of Ecommerce in FMCG', June 2014, http://www.kantarworldpanel.com/global/Publications, (accessed 29 November 2014).

Kantar Worldpanel, Grocery market share data http://www.kantarworldpanel.com/en/grocery-market-share/great-britain/snapshot/03.01.16/18.08.13, (accessed March 2016).

IGD, 'Online Shopping Behaviour Report', September 2011, http://shoppervista.igd.com/Hub.aspx?id=32&tid=4&rptid=189, (accessed 29 November 2014).

Gooding, P., 'Consumer Prices Index and Retail Prices Index: The 2012 Basket of Goods and Services', *The Office for National Statistics*, 2013.

Office for National Statistics, Table A35 – 2014 Detailed Household Expenditure https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/expenditure/datasets/detailedhouseholdexpenditurebycountriesandregionsuktablea35, (accessed 24/09/2017).

Office for National Statistics, 2014 Regional Gross Disposable Household Income (GDHI), https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/bulletins/regionalgrossdisposablehouseholdincomegdhi/2014#regional-nuts1-gdhi-estimates, (accessed 24/09/2017).

Office for National Statistics, Population Estimates Analysis Tool for Mid-2014, https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesanalysistool, (accessed 24 September 2017).

HM Revenue & Customs, 2014 UK Disposable Income Distribution, https://www.gov.uk/government/statistics/percentile-points-from-1-to-99-for-total-income-before-and-after-tax, (accessed 24 September 2017).

Office for National Statistics, Table B5: Characteristics of households in 2014 U.K., https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/compendium/familyspending/2015/familyspending2015referencetables, (accessed 24 September 2017).

Office for National Statistics, Spreadsheet A01: Labour market statistics summary data tables for July-September 2014, November 2014, http://www.ons.gov.uk/ons/datasets-and-tables/index.html, (accessed 29 November 2014).