Strategic Customer Relationship Marketing and Re-intermediation Models in the Insurance Industry

Alun Lloyd Brain

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
For the degree of Doctor of Philosophy

April 2014

DECLARATION

I certify that the substance of this thesis has not been already submitted for any degree and is not currently being submitted for any other degree. I also certify that to the best of my knowledge any assistance received in preparing this thesis, and all sources used, have been acknowledged and referenced in this thesis.

Abstract

Strategic Customer Relationship Marketing and Re-intermediation Models in the Insurance Industry

This research uses a case study of a UK car insurance company to investigate the relationships among price aggregator (re-intermediation) purchase channel, purchasing habits, marketing response models, marketing mix variables, business models, and strategic customer relationship marketing. The introduction of aggregators within the industry has changed the UK car insurance environment substantially in terms of the above core aspects. The research explores the following questions. How do the insights map to the particular business contexts of the case company and its drive for sustained growth and profitability? How does re-intermediation relate to strategic marketing planning and implementation via the marketing mix? How can the results help to reposition the case company with regards to its future growth and profitability through an integrated business model? How has the performance of existing distribution channels been affected by the advent of price comparison models?

A wide range of statistical models and data mining tools were applied to this research, including vector autoregressive (VAR) modelling, general linear regression, quantile regression, autoregressive, moving average; autoregressive integrated moving average, GARCH, logistic regression; decision trees and neural networks models. The research also uses scenario testing for business model understanding and hypothesis testing for marketing framework. These methods allowed the researcher to better understand the new aggregator enriched environment.

By way of main theoretical and practical contributions to knowledge, the study provides an in-depth knowledge of the insurance re-intermediation problem and the construction of an Integrated Business Re-intermediation Model (IBRM) that enhances growth and profitability of company x, and insurance companies in general. This the first to study the effects of reintermedation within the UK car insurance industry which compares the business prospects of the case car insurance company pre- and post-joining an aggregator. The analyses show that price aggregator channel significantly interact with other channels in influencing the customer retention rates and life time values available to the company and hence its future growth and profitability. Insights from the IBRM model could be used to develop the car insurance and related businesses further.

ACKNOWLEDGMENTS

First and foremost, I would like to express my deep gratitude and appreciation to my Director of Studies, Dr Patrick Oseloka Ezepue, for his endless guidance and encouragement. Also to my wife who spent endless hours for her support and has over the past 6 years become knowledgeable in marketing and statistics through frequent discussions. I also extend my grateful thanks to my second supervisors, Dr Kassim Mwitondi and Dr Frances Slack for their generous support. My thanks are also due to Richard Johnson who helped me with time series regression at the beginning, Ben Smythe who introduced me to Quantile regression, Sameer Rahman who provided additional insight on the final IBRM and finally those lovely people who let me use their data.

DEDICATION

This dissertation is dedicated to
My wonderful wife, Juliet
My lovely children: Alex, Eva, Harry and Esme
David T Brain

My Gran, Parents, brothers and sister
AW Jones, Dave Owens, Ed, Gandhi, Jonah, Steve, Tom
My many friends for their support

Contents

Abbreviations

Item	Interpretation
ACR	Aggregator Conversion Rate
AIC	Akaike Information Criterion
ARIMA	Autoregressive Integrated Moving Average
ASE	Average Square Error
BIC	Bayesian Information Criterion
CsTK	Contributions to Knowledge
CRM	Customer Relationship Modelling|Management
DK	Domain Knowledge
DM	Direct Marketing
GARCH	Generalised Autoregressive Conditional Heteroskedasticity
IBRM	Integrated Business Reintermediation Model
IRF	Impulse Response Function
KPSS	Kwiatowski-Phillips-Schmidt-Shin unit root test
L.H.S.	Left Hand Side
MCMC	Markov Chain Monte Carlo
OFT	Office of Fair Trading
PS|W	Problem Studied | Why
RFM	Recency Frequency and Monetary methodology
VAR	Vector Autoregressive
VEC, VECM	Vector Error Correction and Vector Error Correction Model

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Introduction

The purchasing of car insurance has changed dramatically in the last thirty years. The Road Traffic Act 1988 (c. 52) requires that ‘a person must not use a motor vehicle on a road… unless there is in force in relation to the use of the vehicle by that person such a policy of insurance’ (UK statue law data base, no date). This law makes it a legal requirement for drivers in the UK to purchase car insurance and as such car insurance is often viewed as price inelastic: ‘the overall demand for these products does not decline significantly when the price increases’ (Hoyt *et al*., 2006, p.8). This does not mean that the market is not competitive, as customers will tend to go for the cheapest price.

Price comparison sites (aggregators) have had a major effect on both the way people buy their car insurance and the car insurance industry itself (David, 2008). Therefore, the focus of this thesis is to explore the impact price comparison sites have had on the car insurance market.

This research uses an established UK car insurance company as a case study. The empirical investigation of price comparison sites will provide useful information about the effect of price comparison sites on the car insurance industry from purchasing habits, marketing, marketing response models, business models and strategic customer relationship marketing. These issues have been discussed separately in the literature, but not combined for UK car insurance (Morgan *et al*., 2006; Keller and Lehmann, 2006; Papatla and Liu, 2009; Stone and Foss, 2002; and Hanssens *et al*., 2003).

Moreover, this research will have wider implications for market re-intermediation in general, including for example the effects of social networking and electronic media channels and as is the case with music and e-books, and platform migration (use of phones instead of computers, for example). Hence, to make this research contemporary the researcher explores at appropriate sections of this thesis channel conflicts and electronic media distribution channels.

1.2 Rationale for research

The key rationale for this study is as follows. Firstly, there is a dearth of research into the UK car insurance. Secondly, although there are some studies of reintermediation, there is no comprehensive study of the key aspects of how aggregators have affected the UK car insurance industry. Studying these key aspects will provide a better understanding of how price comparison sites have affected the UK car insurance industry. This study explores these aspects using data from an established car insurance company, particularly the business model and strategic customer relationship marketing ideas.

Financial aggregators are a relatively new phenomenon, so this research is the first to measure their impact within the UK car insurance industry, albeit for the case company. The research, therefore, compares the behaviours of the case car insurance company pre- and post-joining an aggregator, as revealed by the key themes studied. Understanding these effects will provide insight for senior managers of the insurance company involved as well as other industries who wish to include aggregators in their distribution mix.

As will be described later in this thesis, the research will enable informed decisions as to whether aggregators will be beneficial for their own companies or not, in any country that has price comparison sites. For instance, companies will understand what main parts of the business are affected for example sales, marketing and customer retention. The emphasis, therefore, is that all the channels through which the customer can make contact and purchase their car insurance are fully explored.

This knowledge will also allow senior management to focus their marketing activities for improved efficiency and operations. Also, this research will allow senior management to focus on not just gathering new customers, but on retaining profitable customers.

By using appropriate statistical and data mining tools to study the full impact of aggregators within the UK car insurance as a case study, the research contributes to the literature base on strategic customer retention analysis and market response modelling.

In summary, this study provides a comprehensive analysis of the effects of price comparison sites on a case company within the UK car insurance industry. Moreover, it provides some insights into the current state of the UK car insurance industry and how companies are adapting to the industry's new reintermediation channel of distribution.

1.3 Rationale for choosing the UK car insurance as a case study

Car insurance is a legal requirement in the UK, so all car owners in the UK must have a car insurance product. In 2011 the UK car insurance industry received £13.3 billion in premiums and insured 23.8 million private vehicles (ABI, 2012). These figures demonstrate the significant size of the industry and the wider business implications of this research.

The research will provide a detailed insight into the UK car insurance company which is used here. This study will help this company understand their efficiencies and any weaknesses in their current strategy. As mentioned earlier, the results have wider marketing implications for electronic commerce-related work in other industries (Doherty and Ellis-Chadwick, 2010).

As a tool for customers to contact an insurance company, price comparison sites represent 'a relatively neutral business model with respect to the buy side' (Domowitz, 2002, p.154). This gives the power to the customer and leaves the car insurance companies to adapt to their new business climate. This demonstrates that customers and insurance companies need to work together so they can both reduce costs and thus be beneficial to each other. Interest therefore lies in this relationship being a reflection of market response within the UK's car insurance industry.

Numerous papers discussing marketing response models, disintermediation, reintermediation (cybermediation) and strategic customer relationship management tend to keep all these subjects separate and do not consider the UK car insurance industry (see, for instance, Bouwman *et al*., 2005; Dumm and Hoyt, 2003; Sophonthummapharn, 2009; Verhoef and Donkers, 2005). Hence, studying these issues together will provide a better understanding of the industry in its current state and provide a further strategic insight for the case company used in this research.

The arrival of price comparison sites has produced a major shock to the distribution channels. Financial price comparison sites are not just located in the UK, but are also being used in other EU and non EU countries, which means that this research could be adapted for business development in other financial organisations and countries.

1.4 Research Issues

As noted earlier, this study aims to assess the effect of aggregators have had for the case company using some statistical models and data mining tools in studying the holistic impact of aggregators on key business aspects such as customer value, retention, relationship management and strategic marketing, which contribute to the growth and profitability of an insurance business.

1.4.1 Objectives of the research

The specific objectives of the research are as follows:

1. To explore the effects of re-intermediation and the marketing mix on the profitability of car insurance business in general, using an established car insurance company as a case study.
2. To explore the suitability of own/different types of business models for car insurance financial management for an established car insurance company, and map alternative scenarios which will guide the future management of growth and profitability of an established car insurance company based on the business models
3. To explore the implications of these results for acquiring and retaining customers in the context of re-intermediation for the case company.

1.4.2 Research questions

The main research questions associated with the research objectives are as follows:

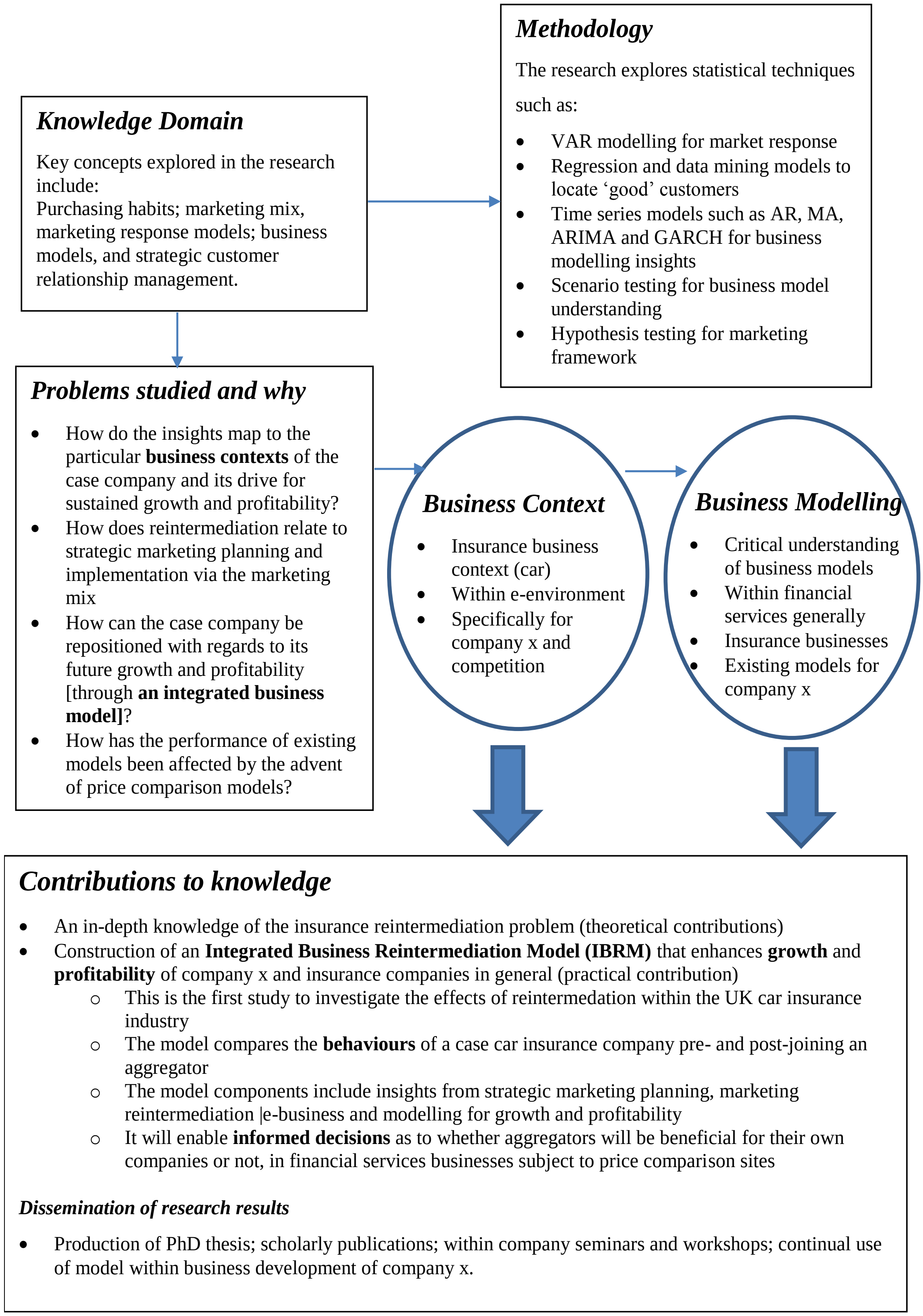
1. ***Research Question 1:*** *How does reintermediation relate to strategic marketing planning and implementation via the marketing mix, in helping the researcher to reposition the case company with regards to its future growth and profitability?*
2. ***Research Question 2:*** *How has the performance of existing channels been affected by the advent of price comparison models with respect to customer retention, new business and profitability?*

Based on insights from the foregoing research questions, it is expected that the study will yield an improved understanding of business process modelling for managing growth and profitability in the car insurance industry in the context of re-intermediation. Given the limited number of similar studies in the field, the research findings may provide useful input into future studies beyond the UK.

To reflect the linkages among the key aspects in the research, including the knowledge domain (DK), problems studied and why (PS|W), methodology and contributions to knowledge (CsTK), the following conceptual framework is provided for this research.

In order to make the chapter more focused on the rationale for the study and the study objectives, further details on the background to the study are presented in Appendix 1 of the thesis.

Figure 1.1: Conceptual framework of research



1.5 Indicative structure of the thesis

This thesis is divided into 8 chapters. Chapter 1 provides the aims and objectives of the study, research questions and the study motivation.

Chapter 2 presents a general literature review of the different aspects of the study related to the objectives and following the broad headings in the conceptual framework.

Chapter 3 introduces the research methodology including the sources of data, the underlying principles and data analysis techniques.

Chapter 4 explores the effects of reintermediation on different customer acquisition channels, acquisition rates, retention rates, and marketing spend. The effects of aggregators are monitored using a vector autoregressive (VAR) modelling approach.

Chapter 5 uses alternative scenarios and different statistical modelling approaches to investigate different aspects of the business model such as marketing, sales, retention and return on investment.

Chapter 6 implements related customer segmentation analyses based on predicted customer value, actual customer value and predicted renewal rate.

Chapter 7 reviews different marketing techniques in order to develop a new marketing framework for the case company (and hence the UK car insurance industry).

Chapter 8 summarises the research findings, including the implications of the research for the UK car insurance industry. This chapter also summarises the main contributions of the research to knowledge

Chapter 2: Literature review

2.1 Introduction

This chapter critically reviews the literature on econometric and applied statistical/data mining models as applied to marketing, customer campaign and response modelling (customer analytics) and reintermediation, with an emphasis on the UK car insurance market. The chapter starts with an overview of the knowledge domain, followed by the theoretical issues studied.

Car insurance in the UK is not a well-researched area; one rare study of the UK car insurance industry was discussed by Blythe and Hackley (2005), who noted how Frizzell Insurance targeted its audience. They found that Frizzell Insura

nce had adapted their marketing so that it attracted one particular type of customer and did not attract high risk customers.

In this section the emphasis is placed intuitively on the key aspects of these issues, whilst more technical treatment of related ideas is presented in appropriate subsequent chapters of the thesis: for example, ideas related to customer relationship management are explained briefly in this chapter and developed in more detail in chapter 6.

In a nutshell, the strategy for literature review used in this thesis is a distributed model in which key concepts are explained in this chapter and more technical aspects of these concepts, including relevant modelling equations are presented in future chapters. For easy follow-through, the chapter headings are informed by the key sections of the conceptual framework (CF) in chapter 1 (figure 1.1) of the thesis.

2.2 Knowledge domain

2.2.1 Purchasing habits

Consumers use the different distribution channels to gather information for shopping. Weltevreden (2007) found that customers will use the internet as a source of information for their purchases from a bricks and mortar shop, and vice versa. Companies that make themselves available on different platforms enjoy the exposure of being more contactable, even though there may be some conflict in the channels.

The way a customer can purchase car insurance in the UK is always evolving, for example from brokers, to phone, to internet and mobile phones. With UK price comparison sites (aggregators) trying to persuade customers to use them instead of going to the insurance companies directly, this is having an effect on the way consumers purchase their car insurance. 'Firm marketing efforts, channel attributes, channel integration, social influence, situational variables, and individual differences' (Neslin *et al*., 2006, pp.101) all contribute to the customer purchasing journey. From the list there is not one overriding factor, so it could be best practice for a company to use as many contact points as possible, especially when aggregators have their own marketing spend.

How people purchase products is very important with the internet opening up new avenues that were previously unavailable. In 2012, price comparison sites in the U.S. are not prevalent as in the UK, with the traditional distribution channels still the main choice for car insurance purchases (Honka, 2012). This could be due to factors as a reluctance to change shopping habits, as well as internet speeds. Another factor to consider is that within Europe, the UK has one of the highest proportions of online shopper (Ofcom 2011). This demonstrates how new financial price comparison sites are and how their impact in other countries is still in development

To predict that all countries will be adapting financial price comparison sites in the very near future would be short-sighted as they require both fast internet speed and for insurance companies to be fully compatible.

Price comparison sites display car insurance quotes a customer would get if they contacted the insurance company directly, so the price strategy of contacting the company directly has to be amended for the price comparison market. Ataman *et al*. (2010) found that the two main factors that affected sales were product and distribution. The product criteria must be relevant and whereas insurance can be bought at three different levels (comprehensive, third party fire and theft and third party only), where comprehensive cover tended to be the most the most preferred choice. Although this study does not consider the different types of insurance cover, this would provide a potential avenue for future research.

The distribution of the product must also be considered in the context of evolving technologies. For the UK car insurance industry, the internet has changed the way people shop and gives customers the choice of going to numerous companies without speaking to anyone to get a price. Price comparison sites have expanded this by giving the customer a choice of getting quotes from many companies by just visiting one website. Atamen (2009) does not mention the importance of how the distributor gets the customer in the first place, and also the product’s brand equity, which is generated through advertising.

2.2.2 Marketing

Selling car insurance directly to the customer is not new, with Direct Line adopting this route in 1985, but this does not mean it should not expect any shocks to the system. When a company does experience a shock, for example the arrival of aggregators, it has been shown that marketing can play ‘an important role in turning around declining performance’ (Pauwels and Hanssens, 2007, p.307). If the company were to keep its advertising the same and remain non-adaptive to its new environment, then this may cause the company to lose market share. Marketing departments need to evaluate the situation thoroughly and in the case of price comparison sites, either join them or fight them. Whichever scenario the company chooses, they must change their marketing strategy.

The main goal of marketing is to attract customers to the business. Marketing has to ‘capture the tastes and standards of every one of its targeted market segments (Meyer and Schwager, 2007, p.10). Marketing can be considered as one of the core departments in any business. It is up to the marketing department to arrange the strategy that attracts the company’s ideal customers, while at the same time staying within the confines of the marketing budget. Marketing provides the face of the company to the public while also needing to be adaptive to new surroundings.

Without marketing, there will be no way of distinguishing one company from another in the same industry so effectively. The arrival of aggregators has meant strategic marketing planning is increasingly important for the car insurance industry. Marketing develops a company’s brand equity, the value of the brand, which companies can increase through marketing communication effectiveness and brand awareness (Stahl *et al*., 2012). Within the internet environment (online) the company’s website presents the first contact between the customer and the insurance company, the ‘customer experience’ (Dayal *et al*., 2000), but aggregators have the ability to influence this experience as they become the first point of contact.

Aggregators can potentially give the insurance company extra customers, so if the company joins an aggregator the marketing strategy needs to be adjusted. Within the marketing department, ‘selling via intermediaries requires that marketing effort is directed at both the intermediaries and the end customer’ (Harrison, 2000, p.91). The insurance company would prefer the customer to contact them directly instead of via an aggregator, as this would save the insurance company money, but with the aggregator marketing budget greater than the insurance company budget, the marketing strategy needs to be modified. The insurance company may have to demonstrate with its marketing that it is a premium product at a reasonable price and that 'customers pay lower prices in aggregate, but not all customers are better off' (Thomas, 2012, p.38). This scenario would hopefully make more customers contact the company directly, but if the customer does purchase via an aggregator, they should expect a high-quality product.

Using the marketing strategy to increase the value of the brand (the brand equity), therefore, cannot be underestimated, either for the companies on price comparison sites or for the sites themselves. It has been shown that 'consumers are likely to be more receptive to trying on-line offerings from a trusted brand name' (Delgado-Ballester and Munuera-Alemán, 2005, p.193). Due to aggregators being online, this does not mean that the traditional (off-line) marketing should not be used. A company that has developed a good reputation off-line can expect the reputation to be transferred to the internet. Building brand trust will help brand equity, which could give an insurance company a better standing when being compared against a different insurance company with a cheaper price.

Being able to measure the effectiveness of advertising is important for companies, as the marketing department has to keep to its budget and use it wisely. It has been shown that marketing does have an impact on sales, though not always immediately (Pauwels and Hanssens, 2007). To fully understand the impacts of aggregators in relation to marketing, market response models need to be developed.

2.2.3 Marketing response models

Marketing response models have been used in the industry for over 40 years, since the creation of the Bass Model in 1969. Marketing models are normally developed to help businesses understand how productive their marketing spend is within the current market. The models provide information for the business to help them generate greater revenue. Models that want to compare themselves with competitors tend to view all marketing spend as one variable against the competitors' spend.

Market response models are flexible enough to be used for different scenarios. Onishi and Manchananda (2013) demonstrate how the new media (blogs) and traditional media (TV) interacted with each other to predict sales. Joshi and Hanssens (2009)'s research show a direct relation between marketing and stock performance on movie launches. Simon and Sullivan (1993) and Srinivasan and Hanssens (2009) find relationships with marketing and brand equity, which also affect a company's stock market prices. This demonstrates response models' flexibility and how they can be implemented within the UK car insurance industry in an aggregator environment.

So far the review has discussed the applications of the market response models without mentioning the models themselves, which tend to be complex. Comparing multifunctional and functional forms, it has been shown 'that the linear and multiplicative approximations are too simplistic for capturing realistically the complexity of the sales response phenomenon' (Pantelidaki and Bunn, 2005, p.518). The models tend to be complex due to the interactions between the variables and lag effects of marketing. Interaction between marketing is not new with Borden (1964) using the term 'marketing mix' since 1949. Market response models are typically multivariate statistical time series with different specifications depending on prior knowledge of the relationships among variables and their time-varying behaviours. They are discussed in some detail in Chapter 4 of the thesis.

Market response models are not simplistic and can be difficult to understand, which can cause a problem with implementation. For any type of response model to be implanted 'depends critically on two characteristics…simplicity and robustness' (Hanssens *et al*., 2005, p.433). There are many different advertising outlets, for example TV, radio and magazine. To address all these in a response model may prove problematic in implementation. For the case company, the marketing department does not come from a statistical background with the capability to formulate many different scenarios, but to not give people the choice of different models in itself may inhibit the implementation.

Another issue to consider when implementing a model is its usability within a company. The model needs to be 'simple, robust, easy to control, adaptive, as complete as possible, and easy to communicate with' (Little, 2004, p.1852). This research will need to address how to provide the model and the best format for the company to use. This raises the issue of whether a trade-off needs to be considered between complexity and usability. In a sense, the approach taken in the research is to build insights from the models intuitively into the proposed IBRM for car insurance in a way that marketing staff can understand them. Ultimately the choice should be with the company on their desired method of using the model and the preferred choice of the company. As mentioned earlier, more technical issues of market response modelling are covered in chapter 4.

2.2.4 Business model

Business models help align company strategy to the changing environments of a business; in this research, the IBRM shows how the case company can remain profitable in an aggregator environment. The need for the development of a new business model extends beyond enabling aggregators in the car insurance environment. Teece (2010) states that a business model "cannot be assessed in the abstract, its suitability can only be determined against a particular business environment and context" (ibid., pp.191). From this point of view, the business model must be specific to its environment and would be difficult to use in a completely different scenario, e.g. selling petrol at a petrol station. The business model applied should assess the way a firm combines a value proposition with supply chain management, the interface with customers, and a revenue (Boons and Lüdeke-Freund, 2013). Additionally, Girota and Netessine (2013) note that whenever new technologies are developed, there is a lack of business models to accommodate these. Thus this research contributes to the development of a business model that encompasses aggregators, by focussing on the generating customer value for the customer as well as for the company.

Between 1995 and 2010, the number of publications referring to 'business model' has increased substantially, but without a unified view of the concept (Zott *et al.*, 2011; Lambert and Davidson, 2013). Within the research carried out by Zott *et al.* (2011), four important themes about business models were discovered: 'a new unit of analysis, a systematic perspective on how to do business, encompassing boundary-spanning activities… and focusing on value creation' (ibid, pp.1038).

Casadesus-Masanell & Ricart, (2010, pp.195) considers the first theme as a 'reflection of the firm's realized strategy'. The car insurance industry is full of competitors, so its unique position will need to be predominantly price-based, while maintaining good customer service to its customers. The second theme is to apprehend how the insurance company can connect its technical marketing and analytical potential with the recognition of financial value (Chesbrough, 2007). For the case insurance company, they would need to consider the costs of implementing an aggregator into its customer contact channels. The third theme can be linked to how they can conduct their business with the customers and partners (Frankenberger *et al.* 2013). Implementing a price aggregator would introduce new partners and components into the whole business. A car insurance company that uses direct avenues only for the customer to contact them will need to evolve to allow third party companies to gather quote information and still be able to reach the customer. Finally, the fourth component concerns value creation: efficiency and novelty (Zott and Amit, 2008). Efficiency value for the company considers cutting costs, and the novelty involves new avenues for conducting business transactions, which is the whole basis of aggregators. For the company used in this study, the value proposition also involves creating a car insurance product that can satisfy the customers' needs at a low price, while still generating profit for the company. These themes are 'interconnecting and mutually reinforcing' (Zott *et al*. 2011 pp.1038) and the following notes use the themes as the basis for the rest of review about business models.

Aggregators will have a profound effect on the car insurance company's business model, even if the insurance company chooses not to integrate them into their distribution channels. Due to time, money or even technological constraints, adapting new customer contact channels can be too difficult for some companies, thus affecting their value to the customer. Another reason why companies may not wish to expand their contact channels may be due to retention rates. Konus *et al*. (2008), Ansari *et al*. (2008) and Neslin *et al*. (2006) note that people who use a choice of multichannel distribution networks tend to be less loyal. Giving the choice of a multi-distribution contact strategy, especially on the internet, makes it easier for customers to compare prices. The easier it is for customers to compare prices the easier it is for them to switch (Israel, 2005). Multi-channel distributions may reduce customer loyalty, but within car insurance the main reason why customers leave their current company is due to their bad claims history, which increases their renewal price (Cohen, 2012). A customer's claims history is important when considering the customer value, as is explained in more detail in chapter 6.

The company has to weigh up the pros and cons of implementing a multi-channel distribution, since not only could this cause conflict within the company and possibly erode loyalty (as it gives customers the chance to search other companies), but also by not enhancing the new distribution technologies, the size of the company may be affected due to customers preferring the new contact channels.

The car insurance business model will need to consider the perceived value of the product, especially when the company is on a price comparison site. The company's perceived value is usually communicated through its marketing, which is discussed in more detail in chapter 7. Sethuramn *et al*. (1999) find that a lower priced brand can be affected when a higher priced brand has discounted its price. If a perceived better quality brand reduces its price and enters a lower price brand, customers would prefer this brand over a consistently low price brand. If the company markets itself as being a good product with high standards with a price that seems reasonable, this will have a strong effect when the price is being compared. As car insurance is price sensitive, the company has to be careful and tactical with its pricing. Shapiro (1983, p.678) notes that ‘high quality items sell for a premium above cost’. In other words, the customer is willing to pay extra for what they perceive as a ‘good name’. With high quality comes a good reputation, which can be perceived positively by the customer.

Regulation, new service and product offers, pressure to reduce costs and customer expectations tend to be the four main factors that can affect the insurance business model (Labush and Winter, 2012). Price comparison sites belong in the new service segment as they change the way the product is sold. The way a company positions itself in relation to other companies may have a profound effect on customers’ choice. How a company differentiates itself from its competitors is known as brand equity or its branding. Brand equity is acquired through marketing the unique selling points of the brand itself: the product; the name; the promotion, and its complete presentation (Murphy, 1992). Branding allows the company to be more recognisable over other companies. ‘A brand is an entity that offers customers (and other parties) added value based on factors over and above its functional performance’ (Knox, 2004, p.106). Price comparison sites need customers and they need to develop their own brand equity, as well as that of the insurance companies they compare.

If an insurance company has robust brand equity, this will also benefit the price comparison site, as the price comparison site can advertise that it has big name brands. Rios and Riquelme (2008) researched online companies and their brand equity, examining brand awareness, value of the brand, trust and loyalty. Their research found that ‘brand loyalty and brand value associations directly create brand equity’ (ibid, p.735), but they also note that customers need to be aware of the company. This shows that the company cannot rely on just the internet for brand equity and that they still need to use traditional advertising methods, which must be incorporated into the car insurance company's business model.

Companies that market themselves as high quality may, however, get away with charging more money. Although this idea has not been fully investigated on aggregators, it provides further opportunities for future work. Blattberg and Wisniewski (1989) and Glynn and Chen (2008) find that higher quality brands get their custom from their own price tier competitors, but lower quality brands rarely get sales from higher tier customers. This demonstrates the power of brand marketing and how perceived value from a brand affects its sales. If, however, a company sets an unrealistically high price – effectively ‘pricing themselves out of the market’, or if a brand of similar quality is sold at a lower price, then this will affect consumers’ view of the company negatively.

2.2.5 Strategic customer relationship

Car insurance companies can grow their customer base by taking customers from their competitors, so retaining their customers is important. Customer retention is therefore important, and has become a major factor as ‘the competitive nature of the insurance industry continues to evolve … and the importance of relationship marketing practices and customer retention continues to grow’ (Taylor 2001, p.32). Increasing customer loyalty practices should be as standard in all industries, along with a clear indication of which customers it would be profitable to retain.

Payne and Frow (2005) note that there is a lack of agreement on the definition of Customer Relationship Management (CRM). There has been discussion whether CRM should be considered as a strategy (Payne and Frow, 2005; Duffy et al 2013; Mukerjee, 2013) or to be part of a general management process such as a system for CRM (Ajmera, 2013; McGrath 2010; Foss *et al*. 2008). Within the car insurance, CRM should be considered both as a strategy and a process, so that this research supports CRM strategy and process as relates to car insurance.

CRM as a strategy can be defined as a:

> *Business strategy and mode of operation deployed to maintain and develop relationships with profitable customers, and manage the cost of doing business with less profitable customers*
>
> Stone and Foss, 2002, p.14

CRM as a system can be defined as a:

> *Technology-based business management tool for developing and leveraging customer knowledge to nurture, maintain, and strengthen profitable relationships with customers.*
>
> Foss *et al.* 2008, pp.69

The implementation of a CRM strategy helps businesses to generate more profit from customers by either cross-selling products or by getting the customer to continue purchasing from the company (retention). The CRM system supports the strategy, since without the system there would be inadequate knowledge of how to contact customers or who the company's most profitable customers are. Using CRM as a system can improve data-driven marketing strategies and provide a holistic profile of the customers across different contact points (Internet, email and telephone) (Even *et al.*, 2010; Verhoef *et al.*, 2010). The availability of high-quality data within a CRM database is high in many applications including car insurance, but developing clear data-driven strategies and procedures to enhance the business performance from numerous CRM systems is still unclear (Zahay et al, 2012). Regardless of this uncertainty, in this research, the CRM strategy is built upon accumulated data to a case insurance company in UK in order to enhance the customer experience.

Customer relationship management and relationship management may be similar but relationship management deals with building relationships with all contacts, whereas CRM is a more focussed model-based approach which deals with profitable customers (Das, 2009). Being more focussed on the customer may mean contacting the customer directly with promotions/offers. Building a relationship with the customer, with the company becoming 'more than just a company' and offering more services, may make the customer less likely to leave. In the UK car insurance industry price is a major factor, so giving the customer more for their money will ensure greater customer retention.

CRM is an important strategy to capitalise on customers that are worth retaining by building strategic relationships with them. CRM uses customer data gathered from its systems to incorporate better strategies and to use the resources in a more efficient way (Ngai et al, 2009). This allows predictive analysis tools to analyse data within the CRM framework with customer retention being the most prevalent concern (Mizaei and Iyer 2014). Within this research the predictive analysis tools are used for churn rate and customer lifetime value.

The CRM in this research is part of the IBRM which also includes the business model. The CRM creates customer segmentation for customer retention strategies from potential high valuable customers, to those customers of a high risk of costing the company money. Every customer that purchases their car insurance via aggregators, costs the car insurance company money. Retaining the correct customers not only builds the customer base more cost-effectively, but also stops them from leaving only to come back again a year later, possibly via an aggregator, which results in additional expenditure. CRM should be a vital strategy for all companies, including car insurance companies.

Sophonthummapharn (2009) researches the adoption of an electronic customer relationship management (e-CRM), but the main theme, that of building a relationship with the most profitable customers, is still prevalent in his research. This demonstrates that as technologies evolve so must the company adapt to the new surroundings. Contacting customers electronically can not only enhance a company's green credentials, by using less paper, but can also monitor the customer's activity via email reporting tools, for example number of emails opened, number of emails clicked. This can benefit the company as those people who do not respond to emails can then be contacted another way.

CRM should be considered as part of the insurance business model and strategy. Improving the customer experience increases loyalty and should generate recommendations from customers to other potential customers (Sweeney and Swait, 2008; Meyer and Schwager, 2007). For the case company, word of mouth recommendations are a valuable tool as the person who recommends the company, will probably remain loyal to the company, and their recommendation will provide free advertising.

Making the CRM part of the business strategy should lead to increasing the employee experience (making the employee enjoy his or her time at work). This has been shown to have a positive experience on customers, which itself has an indirect effect on customer retention (Mosley, 2007). Improving employee morale will have a positive effect on staff turnover. This has the cumulative effect of ensuring staff are more experienced, as they stay with the company longer, and are therefore more confident and capable of dealing with customers, who after contact with staff are satisfied with the level of customer service received. This should be prevalent in car insurance especially when the customer reports a car accident. Some car accidents can be quite traumatic for the customer, so when the customer first reports their accident to the insurance company they need to be comforted and reassured that they are dealing with someone experienced and knowledgeable.

Price comparisons sites like to retain the companies on as an exclusive basis; that is, that type of car insurance company is only available at this web site. It has been shown 'that intermediaries do not like their suppliers to engage in multi-channel distribution' (Coelho and Easingwood, 2008, p.38), especially if the companies adopt a different pricing structure for each individual intermediary. The more exclusive companies an aggregator has, the better placing it has in the market. On the other hand the consequences of using a single channel for the company are that they are limiting their distribution. By using various distribution channels, including intermediaries, a company can use the advertising of the distribution channel to promote its own product and make contact with potential customers that they normally would not reach.

2.3 Technical issues studied

2.3.1 Business context

As mentioned previously, in section 2.2.4, there are four main events that can affect the insurance company: regulation; new service and product offers; pressure to reduce costs and customer expectations (Labush and Winter, 2012). These are specific to the insurance industry because it is a traditional business (Labush and Winter, 2012). The car insurance industry must always be adaptive to its new surroundings as well as keep an eye on managing down the costs. It is for this reason that the company used in this case study uses the marketing budget to pay for the sales attributed to the aggregator.

Aggregators change the car insurance company's distribution channel as customers do not need to contact them directly. Changes to distribution channels have been researched but tend to focus on companies utilising the internet for customers to contact them directly (Huang and Swaminathan, 2009; Pfiel *et al*., 2007; Wolk and Skiera, 2009) or the impact of insurance companies using a direct channel on a brick and mortar intermediary (Bouwman *et al*., 2005; Hoyt *et al*., 2006; Pfeil *et al*., 2008). The research reflects how companies should expect some channel cannibalisation and that pricing strategies should reflect the new distribution channel.

The adoption of the internet as a new direct contact channel for a UK car insurance company may run into issues in a price comparison site arena. UK car insurance customers on the internet have a choice of two avenues to get a quote, either through a price comparison site or via the company directly. The main selling point of price comparison sites is that they can produce many quotes as quickly as a customer going to one company directly. If a customer chooses the price comparison site route only and the insurance company is not on their website, then the company will lose out, but if the company is on website this may lead to cannibalisation. Price comparison sites also show the price a customer would get if they contacted the insurance company directly, so the price strategy of contacting the company directly has to be amended for the price comparison market. More technical issues of the business context are covered in chapter 6.

2.3.2 Reintermediation and strategic market planning

Selling car insurance directly is quite a mature market, which does not mean it should not expect any shocks to the system. As noted earlier, when a company does experience a shock, it has been shown that marketing can play 'an important role in turning around declining performance' (Pauwels and Hanssens, 2007, p.307). If the company were to keep its advertising the same and non-adaptive to its new environment, then this may cause the company to decline in its market share. Marketing departments need to evaluate the situation thoroughly and in the case of price comparison sites, either join them or fight them. Whichever scenario the company chooses they must change their marketing strategy.

Certain customers have particular shopping habits, whereby they continue to purchase their goods in the same manner constantly. Marketing can be used to change a customer's shopping habits or their entrenched buying behaviour, which is when 'people get used to buying certain products through particular intermediaries and have an inbuilt inertia to change' (Wilson *et al*., 2008, p.531). The other side to this argument is to stop customers going to aggregators to change their insurance company. Aggregators need customers to change their insurance each year to generate the most profit. Marketing techniques have the potential to stop customers leaving their current insurance company and from using aggregators to review their renewal prices.

There are many tools available for car insurance companies to contact their customers to enhance/change their purchasing habits. Loots and Gobler (2013) found that a mixture of technologies worked in building relationships with the insurance customer, including SMS on mobile phones. This technology will allow the company to contact their customer at any time the customer has their mobile phone with them with their latest offers. However, contacting the customer by email, direct mail and texting may lead to the customer being bombarded with too many communications. This may have a negative effect, as the customer may opt out of further marketing mailings or even worse cancel their policy. It is vitally important therefore to maintain the correct equilibrium, which is of course what relationship management is concerned with. More technical issues of marketing in an aggregator environment are covered in Chapters 4-7.

2.3.3 Future profits and growth

Customer growth and profitability can be created in two ways for an insurance company, via new customers and retained customers. The way customers shop for their car insurance is different to other insurance types. Customers tend to buy their insurance online rather face to face (Rokach *et al*., 2013). This gives car insurance the capability to grow at a more profitable rate than the other insurance types. Driving more customers to shop on-line bypasses the need for staff to answer phones and also reduces building running costs. In this instance, aggregators should not be considered as a competitor, but a useful tool to help company growth and profitability by maximising the use of the web. Ideally, for the insurance company, it would be more beneficial for the customer to go to the insurance company only, as this would save them aggregator costs.

Other financial institutions may gather extra income from the customer via cross-selling further financial products, (Kaishev *et al.,* 2013), but for the company used in this study, this is not an option, as they deal in car insurance only. To gather further profit and growth for the company, they will need to consider customer retention. Customer segmentation is a common tool in insurance as well as other industries, used to highlight which customers it would be most advantageous to retain. Labusch and Winter (2012) comment on one company that classifies customer A to D. The A-ranked customers are treated as high quality, whereas the D-ranked customers are the ones the company wishes to leave. As mentioned previously, in 2.2.4, customers with a bad claims history can cost the insurance company a significant amount of money. Customer segmentation of profitable customers is detailed further in chapter 6 of the thesis.

If a customer is determined to leave their current car insurance company, they will. As mentioned previously, aggregators have reduced the switching costs so locating potential valuable customer is a necessity. Customer campaign and response modelling are used to maximise profits by either targeting potential customers or targeting profitable customers. Unlike mass media marketing (e.g. TV, outdoor posters, radio), direct marketing, as the name suggests, approaches its clients directly. The most common model used tends to be the Recency, Frequency and Monetary (RFM) framework, mainly due to its simplicity in application. RFM is a segmentation technique that can be applied to any service-based business including car insurance (Nanni *et al*., 2013):

- Recency: time since the customer made his/her most recent purchase
- Frequency: number of purchases this customer made within a designated time period
- Monetary: average purchase amount

RFM has been compared with other techniques to pinpoint the most profitable customer. Olson *et al*. (2009) compare an RFM model with data mining techniques (logistic regression, decision trees and neural networks) to evaluate customer response models and find that data mining techniques provide a more useful tool than RFM in locating the most profitable customers. Data mining can involve a wide range of variables to help build its model and can give individual customers a score, but it also has its faults. Logistic regression results can be complex to understand, decision trees need to keep the number of rules applied small for ease of interpretation, and neural networks work like a black box, with the user not fully understanding how it achieves its results. More technical issues of the data mining are covered in chapter 6. An issue with RFM is that it can have a high correlation between the frequency and monetary metrics, but RFM is much easier to understand and implement, thus its popularity. In this study, these different approaches to customer segmentation are attempted in order to compare their potential merits to the case company.

Another problem with RFM is that it cannot be used to locate potential customers, 'due to there being no information regarding potential customers' (McCarty and Hastak, 2007, p.657). This will raise an issue with the company as sales teams need to be able to contact potential customers and without a profile of the customer available to make sales to, this can be problematic. Typically, sales teams need to use externally purchased data on consumer attributes and behaviours and conduct own surveys of potential customers in order to augment the internal database with these data. Atypical example of such data gathering may include gathering car tax renewal dates, which may fall on car insurance renewal date as well. Another issue with the RFM model for insurance is that higher monetary amounts usually signify a greater risk to the company.

As indicated above, an alternative to RFM is to use data mining tools for marketing response metrics. They tend to work more favourably with direct marketing, as they can help pinpoint the best responders to a certain piece of direct mailing and improve return on investment. Data mining is often used in response modelling which can benefit customer relationship management by suggesting 'the best time to make a cross-sell or up-sell offer' (Berry and Linoff, 2004, p.121). The main problem with data mining is the complexity of the results when they are presented to management personnel. This issue may, in effect, reduce the popularity and therefore the implementation of a data mining strategy. As a result despite the limitations of RFM in comparison to data mining, RFM remains a popular method but, this does not mean that data mining cannot be used with RFM. The application of data mining with RFM is explored further in chapter 6 and the decision to build in their insights into the IBRM framework addresses the problems of complexity and usability on the part of marketing staff.

2.4 Contributions to knowledge

2.4.1 Theoretical contribution

This is the first research to study the effects of reintermedation within the UK car insurance industry. The research applies structural time series models such as Vector Autoregression (VAR) and Vector Error Correction (VEC) models to the analysis of the differential effects of price reintermediation (use of price comparison sites) on different channels and marketing data used in insurance marketing. Whilst the statistical theory of structural time series is well-developed in the literature, the use of such models in studying the persistence effects of different marketing decisions in insurance, especially in the context of reintermediation, has not been noted by the researcher in the literature. Using appropriate data on car insurance from the UK case company, this research applies such models for the first time in the UK car insurance industry.

Related to the use of structural time series models in car insurance marketing, the research also applies relevant data mining models (neural networks, logistic regression, and decision trees) to the analysis of car insurance data in the light of price reintermediation. The data mining models particularly show how relevant data on customer life time values are used to segment customers into different categories which can targeted differently in order to grow a car insurance business profitably. The results provide comparative insights on how price reintermediation effects could be modelled in order segment car insurance customers appropriately, and thereby support marketing decisions on car insurance retention quantitatively. There is no other study known to the research which uses data mining tools in this way in car insurance marketing.

Importantly, the research develops an Integrated Business Re-intermediation Model (IBRM) integrated business model which combines insights from the statistical modelling of the car insurance data for managing car insurance businesses in light of price intermediation. The model provides theoretical support for understanding how changes in marketing, customer retention practices and purchasing behaviours, especially involving internet-based purchases, can be strategically managed. The model can therefore be adapted by other car insurance companies and also other insurance products which are accessible through price comparison sites. In other words, within this new environment the IBRM links a car insurance business model with a CRM framework to create new theoretical understanding of how price aggregation affects different parts of the car insurance business, for example sales, marketing and IT systems. This is a key theoretical contribution of the research to knowledge.

In summary, the theoretical contributions of the research consist of use of appropriate statistical time series and data mining models in studying different aspects of car insurance including the effects of price aggregation on key marketing variables and other channel effects, and the creation of the IBRM for car insurance within a price aggregation environment. The researcher reiterates that this is the first time such quantitative marketing and model development work has been conducted with a link to sales, channel effects, CRM perspectives, and in a way that addresses the growth and profitability prospects of a car insurance business.

2.4.2 Integrated business model that enhances growth and profitability (the IBRM)

Car insurance purchasing tends to be an annual event, so it would be wise to make sure that contacting customers coincides with their date of renewal. Gönül and Hofstede (2006) note that timing is important for contacting customers as 'distribution costs could be reduced' (Gönül and Hofstede, 2006, p.65). Even though this research uses catalogues, timing or when the person is most likely to purchase should be considered. As purchasing car insurance is an annual event in the UK, people may only be in the market for 3 weeks. If the company were to deliver mailing to the customer every month, this could lead to 11 out of 12 mailings going to waste.

For a complete picture of which technique to use for the UK car insurance, a test would need to be carried out of all the different techniques. Rust and Verhoef (2005) use an insurance company for their research, using direct mailing and loyalty magazines and control groups for each media. Their research compared the RFM model against a Markov Chain Monte Carlo (MCMC) hierarchical model, and found the hierarchical 'produced better predictive results in a holdout sample than segment-based approaches' (Rust and Verhoef, 2005, p.486). RFM models, when compared to other techniques, do not tend to fare well, and in the case of insurance, RFM was again proven not to be as accurate at predicting customer behaviour, hence the use of different modelling approaches in this research. More contributions to the business model that enhances growth and profitability are covered in Chapters 5-7 of the thesis.

2.4.3 Consumer behaviour modelling

With the internet being used for information, this also allows customers to leave their reviews of companies. It has been shown that 'negative online consumer reviews have a more powerful impact on product attitude than positive online consumer reviews' (Lee *et al*., 2008, pp.341-352). By opening themselves up to numerous distribution channels, companies may leave themselves more prone to bad reviews, so they need to keep on monitoring their customer feedback. This may require additional resources from the company to monitor feedback from customers on the most popular social networking sites, to control any negative feedback.

Price comparisons sites like to retain the companies on an exclusive basis; that is, that type of car insurance company is only available at this web site. It has been shown 'that intermediaries do not like their suppliers to engage in multi-channel distribution' (Coelho and Easingwood, 2008, p.38), especially if the companies adopt a different pricing structure for each individual intermediary. By using various distribution channels, including intermediaries, a company can use the advertising of the distribution channel to promote its own product and make contact with potential customers that they normally would not reach. More contributions to the customer behaviour pre- and post-joining aggregators are covered in Chapters 4 and 5 of the thesis.

2.4.4 Informed decisions

When a company concentrates its marketing into one contact channel, this means the focus is on the product. This focus can be strengthened as the company is not altering its message to appease different channels simultaneously (e.g. Direct Line insurance will keep on using a direct method only route). This strategy can also make companies 'concentrate on the cheapest channel system for that product' (Coelho and Easingwood, 2008, p.38). This may be beneficial to the company, but unless there is an available budget to get their voice heard over all the other distribution channels this may lead to the company becoming smaller. Paradoxically then, a company with a strong marketing message may suffer financially if it is unwilling to extend its message to encompass different channels.

As mentioned previously, although a company can get more exposure on the internet using an established distribution channel, the distribution channel also needs the company to improve their brand equity and attract more customers. Trachtenberg and Fowler

(2010) report on how Macmillan publishers were able to negotiate their sales price with Amazon, or they would refuse to be part of Amazon. This comes in direct response to the iPad device, which offers a new platform for customers to view their books. With customers being given more choice of different platforms for their purchasing habits, the B2B scenario will also need to evolve to reflect this.

Another issue companies need to consider about the distribution networks is that other devices may be able to copy their functionality. In March 2011 Google bought a price comparison site, BeatThatQuote (Williams, 2011) and since July 2013 has been competing with aggregators for car insurance. This shows that companies must always look at other avenues and the evolving market to keep up to date. Price comparison sites are still new and evolving, but they need to keep an eye on different platforms that may offer the same functionality, but may charge the financial companies less money, thus making the comparison sites themselves redundant. All of these scenarios demonstrate how customers choose how they wish to contact a company. More contributions to the informed decisions for the case car insurance company are covered in Chapters 4-7 of the thesis.

In summary, the above literature review highlights the need for gathering richer datasets (see section 3.2.1) which will enable (insurance) companies to perform further analyses including the effects of branding within an aggregator environment, wider CRM studies which incorporate employee morale and satisfaction, possible cannibalisation effect of using aggregator channels, monitoring social networking sites, and effective use of external and internal data on customer surveys (McCarty and Hastak, 2007), for example.

CHAPTER 3: Data and methodology

3.1 Introduction

This chapter details the research methodology required for the research summarised in the conceptual framework in Chapter 1 and literature review in Chapter 2. Firstly, the following section contains a brief description of the data and selected data analysis software. This is then followed by the research objectives and questions which are recalled in this chapter for easy follow-through of the methodology. The final section summarises the chapter.

3.2 Data and computer programs

3.2.1 Data

The data provided by the case company can be considered as secondary data. 'Secondary data are data that have already been collected for purpose other than the problem at hand' (Malhotra *et al.* 2012, p115). The data used is derived from the data warehouse, which can be considered structured, regularly updated, reputable and trustworthy (Malhotra *et al.* 2012). The use of the secondary data is important as the case company was able to provide data that cannot be effectively duplicated by another researcher, while also saving time and money

The main disadvantage of the data is the lack of price comparison site data, as this belongs to different companies and not the case study. Also, price comparison data contain additional information such as where the insurance company is located, their ranking on the websites and who their competitors are. Additionally, competitors can affect their prices by changing their excesses, which makes it even more difficult to compare the effectiveness of the brand on the aggregators website. Furthermore, the case insurance company only sells car insurance, whereas aggregators sell numerous products. In this scenario, it is difficult to track any brand effects from a company selling other types of insurance product, e.g. house and motorcycle.

The results from the data used can be used to help other industries considering implementing price comparison sites into the acquisition channel mix. The results show how the business model has to change to implement price aggregation and how conducting business 'as normal' will not suffice.

The researcher notes that despite the lack of richer data sets from price comparison sites noted above, other insurance companies can use similar approaches in the research to explore how to grow their companies by appropriately modelling the effects of price comparison sites on their businesses. It is not expected that other companies will obtain the same results as in this research, so that it is the approaches and model can that can be applied to other companies. Given that the research is the first to develop the IBRM, it is also expected that both the case company and other companies will benefit from using the model as a basis for further studies which will help the companies to continue to grow the companies profitably in the future, although with appropriate modifications as more data are accumulated.

Two datasets from an established UK car insurance were provided for this research. The first dataset contained monthly figures (between the dates of Jan 2006 to August 2009) detailing spend, quotes and sales from the different media channels of acquisition to be investigated. These date periods are relevant as they cover the dates before the company joined a price comparison site, and afterwards. The data provided by the company show that until April 2007 the company only functioned as a 'click button' tool to retrieve a quote. That is, a potential customer had to: go to the aggregator; fill in their details; get a list of quotes; click on this company banner which did not have a quote next to it; go to the car insurance company directly, and then complete their details again on the car insurance web site. Between May 2007 and August 2007 the company was in its testing stage to make sure that its systems and infrastructure could manage with this new channel. From September 2007, the company was fully incorporated with the aggregator.

The second data set contained customer specific data between January-May 2010 (inclusive). The data set contained 189,798 rows of data from policies that were due for renewal. The data contained the fields: age group; allowed to contact policy holder; car age; car colour; claimed on insurance; gender; married; marketing cost; marketing source; no claims bonus (years); no claims bonus protected; number of drivers; pay method; renewal year; social grouping; total claims costs; total premium; type of insurance cover; UK region; customer value, and vehicle group.

3.2.2 Selection of software programs

The main statistical software package used for this research is SAS supported by Excel spreadsheet analysis. The justification for the use of SAS is that it is an industry-wide accepted software capable of modelling large data sets. SAS also has an add-in tool for data mining called Enterprise Miner. It is these tools that allow data interrogation and model building.

3.3 Overview of the research methodology by objectives and research questions

3.3.1 Linking the research objectives and questions

Figure 3.1 below visualises the overall methodology for the research discussed in this thesis. It illustrates the links between the research strategy, objectives, questions and the thesis chapters.

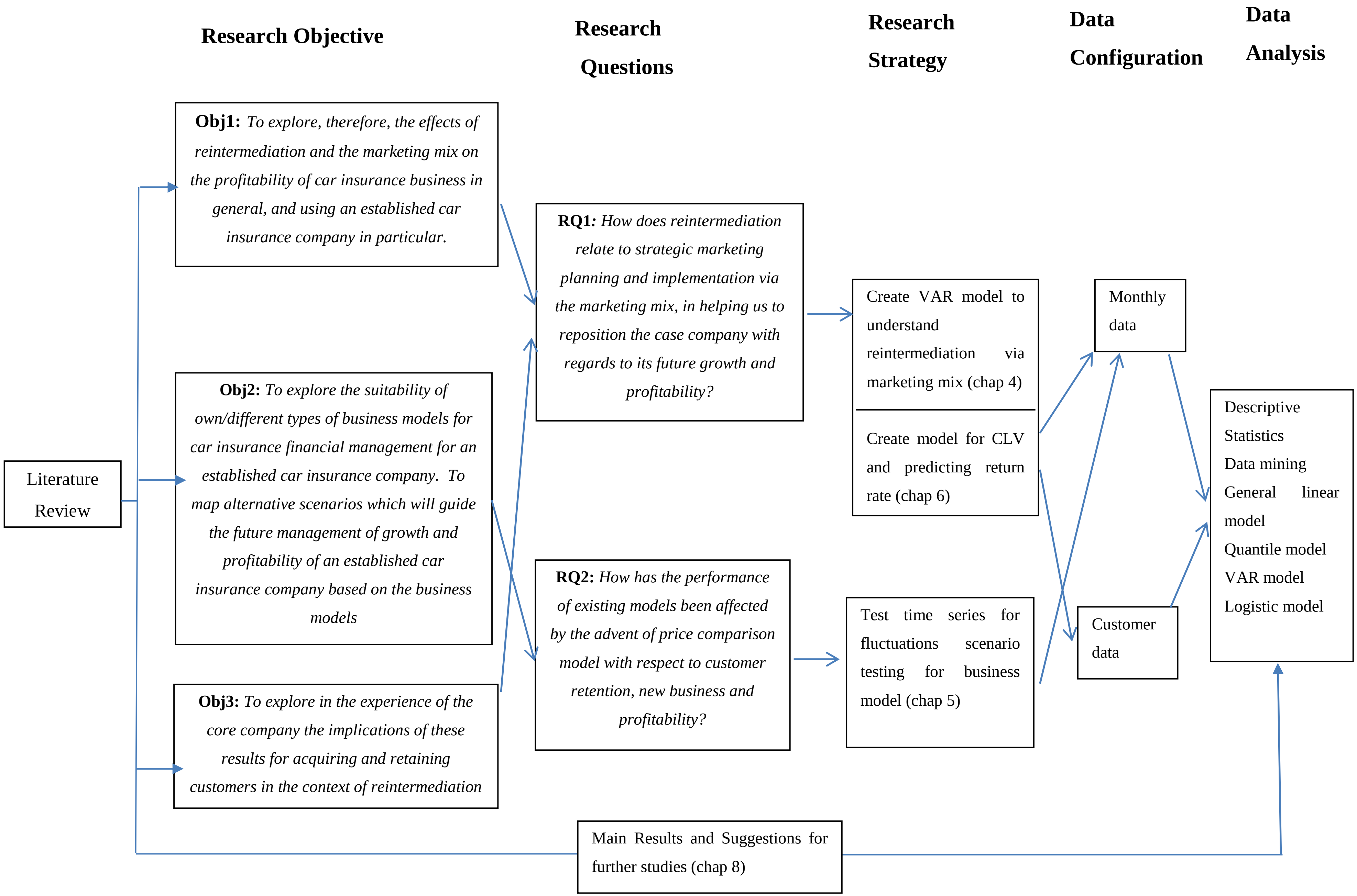


Figure 3.1: Overall methodology for the research with links among the research strategy, objectives, questions and thesis chapters

It should be noted that not all of the data is suitable for all chapters. Hence, the chapters specify what data is relevant for specific aspects of the data analysis.

3.3.2 Summary of the research methodology by objectives and questions

The approaches for investigating specific objectives of the research are as follows:

Objective 1 and 3 (RQ1):

Investigating this objective will require the use of the information gathered from the literature and applying statistical models. The interaction between marketing variables is complex, which is why a vector-autoregressive model will be used to measure these interactions (Chapter 4). Also, by comparing different statistical and data mining techniques, the research can discover the most suitable method to predict customer lifetime value and probability to renew (Chapter 6).

With a recommendation of a method to target the 'right' customer, the research also considers best practices on acquiring new customers. Using the 7P's and hypothesis testing for new marketing framework will be developed (Chapter 7).

Objective 2 (RQ2):

This objective involves a literature review of the car insurance business model with/without price comparison. The research will demonstrate how the introduction of price comparison sites has affected the UK car insurance industry as a whole. The 'what if' scenario will need to cover the quotes, sales, premium and marketing spend aspects of the insurance company. The issues covered would include:

- If the insurance company did not join a price comparison site
- The optimal marketing spend
- If the company chose to abandon a certain media type

The 'what-if' scenarios will need to follow the different stages of the customer journey, from marketing, enquiring about a price (a quote), purchasing car insurance and customer retention. Also for this objective, time series statistical methods are to be used to understand the underlying relationships between marketing and different acquisition channels (Chapter 5).

The statistical models applied in the data analysis and modelling Chapters (4-6) are summarised below.

3.3.3 Summary of methods in key chapters

RQ1 – How reintermediation relate to strategic marketing planning and implementation via the marketing mix?

This research question is spread across three different chapters, 4, 6 and 7. For example, Chapter 4 considers the overall effect of aggregators and channels on each other. Chapter 6 considers targeting specific customers using customer segmentation, while Chapter 7 considers the strategic marketing aspects.

Chapter 4 reviews the concepts and the importance of measuring shocks to a business environment using market response models, for example introduction of price comparison sites into the marketing mix. For this, the Vector Autoregression (VAR) approach to market response modelling is used to examine the direct and indirect effects of customer aggregation on channels and other marketing metrics including the marketing spend, compared to traditional channels, such as TV, radio and the press.

Chapter 6 considers segmenting customers by their value and likelihood to renew. To determine the most predictive tool, decision trees, neural networks, general linear models, quantile regressions and logistic models were constructed for detailed comparison. To limit the influence of outliers, this research also considers Winsorization within the statistical and data mining models for completeness.

Chapter 7 reviews different marketing techniques, to develop a new marketing framework for the UK car insurance industry. The framework would need to consider how different marketing practices affect customer acquisitions and renewal rates.

RQ2 - How has the performance of existing models been affected by the advent of aggregators?

Chapter 5 begins to develop the IBRM business model in order to explore the key aspects of insurance marketing that are affected by the introduction of aggregators. Different time series techniques were constructed: autoregressive; moving average, autoregressive integrated moving average, and GARCH models. These models are applied to different scenarios to establish a deeper understanding of the mechanics within the business model.

3.4 Summary

This chapter describes the overall methodology used in this research by linking it to the research objectives and questions. It provides detail of the data and some of the core themes that require different levels of aggregated data, monthly and at customer level. Also, the chapter summarises the different types of statistical models and data mining tools used in the descriptive analysis of the effect of price comparison environment within a price comparison environment.

Chapter 4: Price comparison and market response modelling in car insurance

4.1 Introduction

The introduction of financial price comparison sites (aggregators) has had a major effect within the car insurance company. By looking at different customer acquisition channels, acquisition rates, retention rates, and marketing spend, the effects of aggregators was monitored by using a vector autoregressive (VAR) modelling approach. An application within the car insurance industry reveals that aggregators have a long lasting positive effect on future acquisition. Comparative results on other market response factors which potentially influence customer acquisition and business performance were explored in the chapter. Hence, the chapter uses the VAR approach to market response modelling to examine the direct and indirect effects of customer aggregation on the marketing spend, compared to traditional channels, such as TV, radio and the press. The results have strategic marketing implications for the UK car insurance industry, by way of optimal allocation of marketing spends on the different customer acquisition channels.

The purchasing of car insurance has changed dramatically in the last 30 years. The Road Traffic Act 1988 (c. 52), requires that 'a person must not use a motor vehicle on a road…unless there is in force in relation to the use of the vehicle by that person such a policy of insurance' (UK statue law database, no date). This law makes it a legal requirement for drivers in the UK to purchase car insurance and as such car insurance is often viewed as price inelastic, so that 'the overall demand for these products does not decline significantly when the price increases' (Hoyt *et al.,* 2006, pp.8). This does not mean that the market is not competitive, as with all purchased products/services, price is important.

The way a customer can contact a UK car insurance company is always evolving. Direct channels involve phones and internet and indirect channels include brokers and price comparison sites (aggregators). Aggregators are relatively new phenomena which operate by comparing many car insurance companies for their car insurance quotes from the information provided by the customer. Aggregators have had 'a major effect on both the way people buy their car insurance and the car insurance industry itself' (David, 2008**).**

How a company operates in a new distribution channel environment will have an impact on old established as well as new companies. Although an established company will have built up its brand equity which will enable it to attract customers more easily compared to newer entrants, some customers will shop on price alone, which may give new entrants a quick start for less marketing. Keller (1993, p.19) noticed that ‘marketing activity can potentially enhance or maintain consumers’ awareness of the brand or the favourability, strength and uniqueness.’ Research into brand equity and customer choice has shown that ‘consumers are likely to be more receptive to trying on-line offerings from a trusted brand name’ (Delgado-Ballester and Munuera-Alemán, 2005, pp.193). A company that has developed a good reputation off-line can expect the reputation to be transferred to the internet. Building brand trust will help brand equity, which could give an insurance company a better standing when being compared against a different insurance company with a cheaper price.

A review of marketing literature shows that the effects of online price aggregation have not been studied in the car insurance industry, compared to many studies of other effects, for example brand equity (Dekimpe and Hanssens, 1995). Indeed, given the competition among insurance companies, it is important that management optimizes its marketing decisions in order to ensure sustained business performance. For this reason, management needs to understand how such marketing decision factors as customer acquisition channels, price, marketing spend, acquisition costs, and retention, interactively influence the profitability of the companies over time. In modelling these effects in this study, the focus is on the effect of the aggregator channel of customer acquisition on business performance indicators, for example acquisition costs, marketing spend and retention.

As mentioned above, the rationale for this study is that there is a dearth of such quantitative modelling in the car insurance industry particularly targeted on understanding the effects of aggregators on marketing performance. To the researcher’s knowledge, this is the first study to explore these effects in the UK car insurance industry.

Modelling the effect of marketing decisions and price aggregation on business performance variables such as marketing spend, customer retention and acquisition costs, is important for a number of reasons. First, it explores the effect of aggregators on customer acquisition. Second, the approach examines the direct and indirect effects of customer aggregation on the insurance company’s marketing spend, compared to traditional channels, such as TV, radio and the press. Thirdly, the results have strategic marketing implications for the UK car insurance industry, by way of optimal allocation of marketing spends on the different customer acquisition channels.

The remainder of this chapter is organised as follows. Section 4.2 reviews the general literature on marketing response variables for example brand equity and distribution channels, and persistence (market response) modelling. Section 4.3 discusses the theoretical background to the statistical models used in implementing marker response modelling and justifies the choice of the VAR methodology for this study. Section 4.4 presents the empirical analysis and the interpretation of the modelling results, with implications for marketing action. Section 4.5 summarises the main results and concludes the chapter.

4.2 Literature review and theoretical framework

As hinted in the introduction, customer acquisition in car insurance industry includes direct purchase of car insurance from companies through telephone and internet, indirect purchase through intermediaries (broker) and more recently through aggregators such as price comparison sites. Direct Line is one of the first companies to start selling car insurance directly to the customer in 1985, thus removing the intermediation role of insurance brokers (intermediaries) from customer acquisition, (a process referred to as disintermediation).

The growth of online marketing in different industry sectors has introduced shock effects on the business strategies and performance indicators of companies in different industry sectors, including mature companies. Examples include the impact of Amazon as an online aggregator in the book market, impact of Netflix in the film industry, and online shopping/home delivery of supermarket groceries by the likes of Tesco, Sainsbury’s and Asda.

Similarly in the car insurance industry, online reintermediation through aggregators introduces shocks in the marketing of car insurance policies which have not been closely studied, hence this UK case study. These shock effects extend to mature channels of car insurance business e.g. direct marketing. As noted above, this research explores similar effects in light of key marketing preforming dynamics and indicators such as impacts of different channels on acquisition rates marketing spend, customer retention and win-back.

When a company experiences a shock, it has been shown that marketing can play ‘an important role in turning around declining performance’ of the company (Pauwels and Hanssens, 2007, p.307). If the company were to keep its advertising the same and remain non-adaptive to its new environment, then this could cause the company’s market share to decline and/or reduce its profitability. Marketing departments therefore need to evaluate the impact of aggregators on market performance and adapt their marketing strategies accordingly. For example, a company may choose to join the aggregators or fight them. Whichever scenario the company chooses, they must change their marketing strategy.

In support of the above points, marketing decisions affect the different channels through which the customers make contact with the company. For example, if a company informs customers that they could save 10% on-line, this will drive them to use the internet channel. It is noted that within marketing there are three channel types: communication; transaction; and distribution channels (Peterson *et al*., 1997). Communication channels include TV, door-drops, and advertisements; transaction channels include telephone and online purchases; and distribution channels include purchases through telephone, web sites, and in some cases intermediaries.

Measuring the effectiveness of marketing within a company is not new, with marketing response models being a common tool to monitor such effects. Marketing response models are statistical frameworks for monitoring customer behaviour and responses in light of business performance variables such as marketing spend, retention, profitability and sales. Majority of market response studies have focused mainly on the direct customer channels which have been in existence longer compared to the online channels.

Even so, the studies have also concentrated attention on the non-insurance markets such as effects of print versus TV/radio advertising on sales of home improvement products (Dekimpe and Hanssens, 1995), pharmaceuticals and fast growing products (Dekimpe and Hanssens, 1999) and myriad studies of the marketing-finance interface (Dekimpe and Hanssens, 2000). The above studies show that as well as the effects of different marketing channels on marketing spend and related business performance variables (acquisition and retention), market response models explore the interaction effects of these channels as single factors and each other. Therefore, as explained in more detail in section 4.3, this study examines the interaction effects on car insurance marketing with a special focus on the aggregator channel. In the researcher's view, this study will be the first market response modelling into the effects of online price aggregation or re-intermediation in the (UK) car insurance industry.

The rest of the literature review discusses the key input variables into the marketing response model, such as word-of-mouth, price comparison sites, retention and win-back.

4.2.1 Word-of-mouth (WOM)

Word-of-mouth is the marketing term for the influence that other people's stories of their experiences with purchases of services/products have on their social group. Before the advent of the internet, this was mainly limited to relatives and friends talking to each other about their experiences, but is now transmitted through the social media sites. Social network sites allow more people to influence prospective customers in their social circle based on their experiences, even if network relationship are not that strong (Lampe, *et al*., 2007). In addition, social network sites can be a low-cost channel to keep customers up to date with a company's communications and enable many visitors to the sites to voice their customer experiences about the company, which influence far more prospective customers than traditional word of mouth (WOM) can achieve.

The use of social networks sites as a WOM channel has been researched previously with an application within marketing. Trusov *et al.* (2009) use impulse response functions derived from Vector Autoregressive (VAR) models and find that WOM referrals have a stronger acquisition impact, compared to traditional marketing and media appearances, on new customers. VAR models are statistical models used for market response analysis which will be investigated in section 4.3. Villanueva *et al*. (2008) explored word-of-mouth (WOM) acquisitions, for example friends and websites and found WOM will double the number of acquisitions compared to traditional marketing.

This differential channel effect of WOM compared to other channels is the kind of insight that market response modelling provides in this line of research. Consequently, the analysis in section 4.4 of this chapter will focus on the effects of WOM aggregators, direct channel on customer acquisition and retention and marketing spend. Lampe *et al*. (2007), Trusov *et al.* (2009) and Villanueva *et al*. (2008) demonstrate how important personal referrals are, whether from social network sites or from other avenues, since they provide a cheap accessible way to communicate to their audience how the market views the company.

4.2.2 Price comparison sites

In the early stages of research into online price comparison sites, such aggregators were thought to be time consuming such that they 'may not be worth much given the small differences in price between different vendors' (Li *et al*., 1999, online). Later research offers a different view with Gorman & Brannon (2008, p.60) noting that 'buyers benefit both in the short run through search cost savings, and in the long run through faster equilibration process to lower transacted prices'. These remarks suggest the need for more studies into the effects of online price comparison sites on customer behaviours and business in different industry sectors, particularly car insurance which this research addresses.

Indeed, it is likely that insurance customers acquired through aggregator channels will behave differently from other customers because of a number of reasons. Aggregators have their own marketing budget to stop customers going to insurance companies directly and it has been shown that 'more price-sensitive customers will gravitate to channels with lower search costs and higher price comparison capabilities' (Granados *et al*., 2011). The lower search cost tends to highlight that consumers are less likely to contact multiple insurance companies, when they can just visit one price comparison site. Also, through aggregators customers may compare companies more regularly on just price alone, especially among the companies which meet key requirements of the customers.

The previous sections clarifies the need of a study of price aggregation effects in the car insurance industry. Consequently, the empirical analyses in this chapter and thesis could be interpreted in light of their implications for marketing decisions by management of aggregator and insurance companies.

4.2.3 Retention

Retaining customers that can generate the most value is an important business goal for any company. Pauwels and Neslin (2008) found that adding a new customer channel affects customer retention. Also, Yoo and Hanssens (2008) researched customer equity based on a single product using six vector endogenous variables, one of which was customer retention. Among other findings, the authors notice a channel effect on customer acquisition and retention behaviours. This shows the need to investigate the differential impacts of different channels when considering market response models, as is the approach in this study.

Given the centrality of retention to customer value which underpins business performance, it is not surprising that customer retention features in most market-based analysis of such performance. A review of the literature on market analyses shows that such customer value metrics as retention are used as multivariate inputs into statistical models that enable marketing modellers to measure the effects of marketing decisions on these metrics, in addition to other marketing and profit-related variables, for example advertising, promotions, own and competitor prices and price differentials, sales calls, and break-even margins (Dekimpe and Hanssens, 1995, 1999 and 2000).

Other studies listed in Dekimpe and Hanssens (2000) model such input/output variables as economic value (profits and brand equity), and customer value (customer intimacy, retention, and acquisition). Shafer *et al*., (2005, p. 201) list the following variables as key components of a business model which different authors examined in the period 1998-2003 - customer targeting, relationships and benefits, revenue and pricing, sustainability, and transaction costs. Winer (2001, pp. 90-91) notes that the relative improvements in business value of the studied companies, which are attributable to 10% improvements in the customer attraction, conversion, and retention factors are respectively 0.7%-3.1%, 0.8%-4.6%, and 5.8%-9.5%. This shows that customer retention is an especially important sub-variable of the customer value component of a business model, and accounts for why analyses of the marketing-finance interface uses retention as an input variable, where appropriate.

The above studies also show, as we shall see later in Section 4.3 of this chapter and Chapter 7 of the thesis, that market response modelling is a form of applied marketing research which uses the marketing mix to determine strategies for improving business value of companies. The sense in which the marketing mix concept is used in this chapter is similar to these studies, which are empirical applications of the standard 7Ps of marketing, which is the theoretical marketing mix with the elements product, price, place (distribution), promotion, people, processes, and physical evidence. Clearly, this research is focused more on place (as with customer acquisition channels) and promotion (as with marketing spend and acquisition costs).

In further support of the above points, other marketing response analyses which explore the cause-and-effect relationships in the customer relationship management literature use variables which are strongly related to customer value, for example improved customer loyalty, customer acquisition and retention, decreased customer costs, and increased profits (Kim *et al*., 2005, pp. 8-9) and pricing in order to achieve profitability (Smith *et al*., 2000, p. 539).

The above points justify why this study focuses on similar variables that capture the customer-economic value potential of marketing decisions in the context of online price comparisons, for example customer retention, win back and marketing spend. These ideas are revisited more technically in Section 4.3 of this chapter, which looks at the statistical underpinnings of market response models, and in Chapter 5 of the thesis which focus on business modelling.

4.2.4 Win-back

Win-back refers to marketing practices used by companies to bring back customer who have left them. Within the car insurance industry, it has been found that winning back customers is more cost effective than acquiring new customer (Evans, 2002). Gee *et al*. (2008, pp. 370) note that 'previous customers are less costly to win-back compared to the costs of acquiring of new customers'; win-back costs can sometimes be cheaper than half the costs of new customer acquisitions; see also Thomas *et al*., 2004). A plausible explanation for this point is the fact that a company does not have to spend more in a win-back scenario, since it has already attracted the customer before and may only need to address particular complaints which made the customer leave. In other words, a customer under a win-back scenario will have experience of the company and thus would have heard of the company before.

In a nutshell, the studies show that win-back customers thus tend to behave differently than retained or newly acquired customers and that having win-back customers potentially improves the effectiveness of marketing decisions significantly. Since these findings are not based entirely within the insurance sector, it is important to explore the win-back scenario in the car insurance industry; hence, the examination of the effect of price aggregation on win-back in this chapter.

4.2.5 Persistence modelling: using market response models to explore long term effects of marketing decisions

An important objective of market response analyses which has implications for the design of marketing strategies that deliver sustainable competitive advantage is finding out how long-lasting the effect of marketing decisions on business performance are (Dekimpe and Hanssens, 1995, 1999). Villanueva *et al*. (2008) note that word of mouth marketing will acquire double the number of acquisitions than traditional marketing, but the effect is long term rather than short term.

Dekimpe and Hanssens (1995) study the persistence of marketing effects on sales using multivariate time-series models of sales and marketing expenditures. These models enable them to determine whether sales are ‘stable or evolving (trending) over time’ and if observed evolution is associated with persistence effects of advertising. Continuing this line of research, Dekimpe and Hanssens (1999) use persistence modelling to explore long-term marketing profitability across four highly insightful strategic scenarios determined by temporary versus permanent marketing effort and response. These strategic scenarios are:

- *business as usual*, for which shocks in marketing effort have short-term impact after which the business performance reverts to pre-intervention levels;
- *escalation of marketing mix activities* without compensating long-run persistence in business performance;
- *hysteresis* whereby 'temporary marketing action causes sustained sales changes'; and
- *evolving business practice* in which 'sustained marketing effort leads to persistent [business performance] results'. The authors provide real-world illustrations and explanations of these scenarios which inform similar applications in different industry sectors.

The strategic choices provided by the scenarios are such that in the case of *business as usual*, a company may achieve marketing profitability by timing and harvesting marketing actions in short enough periods before the business impact peter out, and repeating such actions in the future at carefully chosen time points. With *escalation*, a company should again escalate marketing action in the period when the business impact outweighs the marketing costs. With *hysteresis*, a company typically realizes sustained business value and should identify the winning short-term marketing actions. Finally, in the case of *evolving business practice*, a company should identify and sustain the profitable marketing actions.

In this study, the above mentioned importance of persistence modelling motivates the analyses of over-time effects of customer acquisition channels and marketing spend on selected business performance variables, for example customer win back, retention and acquisition costs, subject to the limitations of the available case data. Section 4.3 of this chapter takes a closer look at the species of multivariate time series-based market response models which support this line of analyses, namely vector autoregression (VAR) models and vector error correction (VEC) models. These models are respectively suitable for modelling *stationary* versus *evolving* marketing and financial time series such as used in this study.

4.2.6 Usefulness of market response models and their strategic marketing implications

Related to the above points (word of mouth, price comparison sites, retention and win back) on the strategic marketing choices available to companies based on modelling scenarios and insights, we note that market response models can be difficult to understand for non-statistically literate company staff, which can cause implementation problems. For instance, there are many different advertising outlets in this study such as TV, radio, and magazines. To address all these in a response model may prove problematic in implementation. Simplicity and robustness are two essential usability characteristics of such models (Hanssens *et al.,* 2005, pp. 433). As further noted by Little (2004, pp.1852), the model needs to be 'simple, robust, easy to control, adaptive, as complete as possible, and easy to communicate with'. These points raise the issue of whether a trade-off needs to be considered between complexity and usability. This research will address this trade-off for the benefit of effective marketing decisions in a car insurance business.

4.3 Theoretical background on market response modelling

This section provides a theoretical background on the types of multivariate time series models typically used in modelling the marketing-financial performance interface as foreshadowed in the above notes on persistence modelling. These VAR (Vector Autoregressive) and VEC (Vector Error Correction) models extend standard univariate autoregressive models, by juxtaposing a simultaneous set of such models in which all the input variables depend on each other according to specified intensity coefficients.

Aspects of the models of interest and their (marketing) applications include: their formal specifications and what the model parameters measure; when and how they are used, for example VAR models for stable (or stationary) time series and VEC models for non-stationary (evolving or trending) series; and key steps in their diagnostics and applications such as statistical tests of model stability and suitable time lags. These aspects are explained in this theoretical background under the subsections 'Model specifications' and 'Empirical Analyses'. Before expatiating on the model specifications, the following notes brief introduce the measurement of some key variables used in the chapter and enable the researcher and reader to make a connection between ideas in the model specifications and the variables.

4.3.1 The method taken to model the effects of price comparison sites within the UK car insurance environment

Using an econometric time series model, the dynamic relationship between the marketing mix and the introduction of price comparison sites can be explored via a vector autoregressive (VAR) model (Dekimpe and Hanssens, 1999). Firstly, the different *endogenous variables* need to be considered as reviewed above. The endogenous variables can be considered as the main effect a customer encounters that prompts them to get a quote from the company. Secondly, the various tests to develop the appropriate VAR model needs to be conducted. Finally, the short and long-term impact of the introduction of aggregators is investigated.

The acquisition rate can be calculated by the total number of car insurance purchases divided by the total number of people who received a car insurance quote within a given time period. For car insurance, consumers may not purchase the same day as getting a quote, which could lead to numerous dates being available. In this instance, to calculate the acquisition rate the researcher will be use the original date the prospecting client first enquired.

$$\therefore\ acq = \frac{N^P}{N^Q} \tag{4.1}$$

where N^P is the number of purchases and N^Q is the number of quotes (enquiries) during a particular period of time. The acquisition rate equation (4.1) will be amended for aggregator, win-back, word-of-mouth and other direct channels, but the same logic applies to all four.

To calculate the retention rate, the channel of purchase will not be considered. Retention rate will be calculated using the total number of people who renewed their car insurance divided by the total number of customers' who were due for their renewal during a given time period. As renewal date would be the same date as the renewal purchase date, there will be no date confusion.

$$\therefore RET = \frac{N^R}{N^R + N^{NR}} = \frac{N^R}{N^T} \tag{4.2}$$

Here, RET is the retention rate, N^R is the number of customer who renewed their policies, N^{NR} is the number of customer who did renew and N^T is the total number of customers who were due for renewal within a given period of time. Since the ratios are valued between 0 and 1, they are transformed using a logit operator, for example for retention

$$RET_t = \ln \frac{ret_t}{1 - ret_t} \tag{4.3}$$

The second stage involved carrying out different tests to develop the VAR model, including

- The suitable *lag order*
- *Stationarity* tests for each endogenous variable, and
- *Cointegration* tests.

The first step in the model approach is to find the lag order which is calculated using Akaike's information criterion (AIC) and Bayesian information criterion (BIC). The AIC test will show the lag order required for the VAR model and the BIC will not only back-up the AIC test results, but will also test whether a Vector Autoregressive Mean Average (VARMA) model will need to be developed.

When modeling a time series scenario, if the mean and the variance of the underlying process are stationary, we can presume that the model will just require a time lag, but in many time series scenarios, the variables tend to be non-stationary. To test for stationarity, Dickey-Fuller unit root test that Enders (1995) proposes and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (Kwiatkowski, 1992) tests will be carried out. The KPSS tests are carried out to verify the Dickey-Fuller results.

If two or more time series, to a limited degree, share a certain type of behavior they can be thought of as cointegrated. Murray (1994) expanded the drunkards walk to explain a non-stationary process, which as a rule tend to be unpredictable, to encompass cointegration. The scenario entails a drunk with an addition of a puppy walking home from a pub.

> *If one were to follow either the drunk or her dog, one would still find them wandering seemingly aimlessly in the night… [but] if you find her, the dog is unlikely to be very far away. If this is right, then the distance between the two paths is stationary and the walks of the woman and her dog are said to be cointegrated of order zero*
>
> Murray (1994), pp.37

The existence of cointegration implies that Granger causality (Granger, 1969; Hanssens *et al*., 2001) exists between the variables, thus making the *endogenous variables* not independent. If a variable can be predicted more accurately using the histories of two or more variable histories, rather than itself, then this means that variables exhibit a Granger-causality. If cointegration exists, then error correction terms are added to correct the cointegrating variables. To test for cointegration, Johansen (1988, 1995a) and Johansen and Juselius (1990) proposed the cointegration rank test by using the reduced rank regression.

By completing the tests for *stationarity* and *endogeneity*, the Vector Error Correction model (VECM) was developed in equation 4.5 (Yoo and Hanssens D, 2005). In this model, retention rate, word of mouth rate, aggregator rate, win back rate, their acquisition rate and marketing spend are the endogenous variables, as they can be explained by their own history and the history of the other endogenous variables (Dekimpe and Hanssens, 1999). In the VAR model, *exogenous variables* need to be included as they affect the model without being affected by it. Three variables which may be included in the model are: time when the insurance joined first price comparison site; seasonal adjustment; and premium/price index.

$$price\, change = \frac{current\, month\, avereage\, price}{previous\, month\, average\, price} \tag{4.4}$$

$$\Delta Y_t = c + \sum_i^k \Phi_i \Delta Y_{t-i} + \Psi Z_t + \Lambda X_{t-1} + \varepsilon_t \tag{4.5}$$

where $Y_t = (RET_t, WOM_t, AGG_t, WB_t, ORAT_t, M_t)$ is a vector of endogenous variables; RET_t is the Retention rate; WOM_t is the Word of mouth rate; AGG_t: is the Aggregator rate; WB_t is the Win back rate; $ORAT_t$ is the Other acquisition rate; M_t is the Marketing spend. $Z_t = (JA_t, JAX_t, PI_t)$, the exogenous variables; JA_t is the dates joined aggregator; JAX_t is the seasonal adjustment variable and PI_t is the Price change Index. X_{t-1} represents the *cointegrating* error terms ($e_{t-1}^{RET}, e_{t-1}^{WOM}, e_{t-1}^{AGG}, e_{t-1}^{WB}, e_{t-1}^{ORAT}, e_{t-1}^{M}$); c is a vector of intercepts; Φ and Ψ are coefficient matrices; k is the lag order and ε_t is a vector of white noise processes with a zero mean and covariance matrix $\mathbf{\Sigma}$. The above concepts, equations and notations are clarified in Section 4.3.3 below.

To capture the dynamic interactions among endogenous variables the coefficient matrix Φ is used. The coefficient matrix Φ only contains the lagged effects of the variables within the VAR system, so existing relationships are acquired by placing restrictions on the residual covariance matrix $\mathbf{\Sigma}$.

The next stage in the modelling involves using impulse response functions (IRFs) that are developed from the VECM Model. IRF are shown to be 'invariant to the reordering of the variables in the VAR [model]' (Pesaran and Shin, 1998, pp.20). Dekimpe and Hanssens (1999) used the IRF derived from VECMs, as the VAR model parameters are not decipherable by themselves (Sims 1980). The IRF will be used to capture the shocks of the introduction of the price comparison sites. Again, the notes in Section 4.5.1 explain in more detail the meaning of IRFs.

The final stage in persistence modeling uses the chosen model to estimate the short and long-term effects of aggregators. From these results, the implications are concluded.

4.3.2 Model specifications

Marketing response modelling enables a company to derive the long-run (input and output) effects of marketing actions by a) capturing the complex interactions of different factor effects and b) interpreting what the short-run effects mean for long-run business performance (Dekimpe and Hanssens, 1999, p. 402). In this research, the marketing variables of interest have been presented above as $RET_t, WOM_t, AGG_t, WB_t, ORAT_t, M_t$.

This section describes the statistical models, Vector Autoregressive (VAR) and Vector Error Correction (VEC) models, which are typically used in modelling the persistence (long-run) effects of marketing *interventions*, for example an insurance company's decision to join and aggregator (AGG) or reduce the marketing spend. The key business performance variables such as retention (RET), marketing spend or profitability (if data on this are available) can be considered as output variables in the models, while the other variables which are known to influence them are considered as input variables. Hence, such statistical models are potentially useful for this study.

VAR and VEC models are species of multivariate multi-equation models in which each equation relates one of the variables in the system as autoregressive and/ or mixed autoregressive moving average regression functions. The idea is that every variable is potentially capable of influencing other variables, and the direction of influence or causality may or may not be known from marketing or economic theory, for example. In this study, for instance, RET and WOM may co-influence one another since this year's WOM could impact next year's retention, and last year's retention may impact the base number of customers who could tell other customers of their experience and thereby impact future WOM effect. In a basic VAR model, because of the cross-variable effects, all the variables are treated symmetrically as joint influencers without the need to specify which are outputs (dependent) or inputs (independent) (Enders, 2010, p. 272).

In a nutshell, there is provision in the character of VAR and VEC models for incorporating prior knowledge or theoretical assumptions in the modelling. The next section presents the models in some detail in order to explain the key model features which justify the choice of model forms and tests used in this chapter.

4.3.3 Multi-equation times series and intervention analysis

Intervention analysis uses a time series to explore short- and long-run impacts of a business decision through changes in the mean of a tine series. Consider the simple example of an autoregressive tine series used in Enders (2010) as follows:

$$y_t = a_0 + a_1 y_{t-1} + c_0 z_t + \varepsilon_t,\ |a_1| < 1. \qquad (4.6)$$

Here, z_t is a dummy variable which takes values 0 for periods t before an intervention and 1 for periods after an intervention (which in that example is the introduction of metal detector technology on the number of skyjacking incidents in the US from the first quarter of 1973 onwards) and ε_t is a white noise disturbance or innovation term with a normal distribution $N(0,\sigma_\varepsilon^2)$. The model is autoregressive because successive y-values depend on previous y-values, which applies to the variables in this study.

Using the lag operator $L^i y_t = y_{t-i}$, $y_{t-1} = Ly_t$ so that (equation 4.6) is rewritten as

$$(1-a_1L)y_t = a_0 + c_0 z_t + \varepsilon_t \text{ which gives } y_t = (a_0 + c_0 z_t + \varepsilon_t)/(1-a_1L) \qquad (4.7)$$

From the properties of lag operators (Enders 2010, pp. 39-42) it is known that

$$(1-a_1L)^{-1} = 1 + a_1L + a_1^2L^2 + a_1^3L^3 + ... = \sum_{i=0}^{\infty} a_1^i L^i \qquad (4.8)$$

so that using $L^i a_o = a_0$ (lags of constants are constants) and sum to infinity of a convergent geometric series $\sum_{i=0}^{\infty} a_1^1 = 1/(1-a_1)$

$$y_t = a_0/(1-a_1) + c_0 \sum_{i=0}^{\infty} a_1^i z_{t-i} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i} \qquad (4.9)$$

Equation 4.7 is an impulse response function (IRF) which enables us to measure the over-time response of y to the intervention z or the effect of z on y. From equation 4.6, the immediate (contemporaneous) effect of z(t) on y(t) is $dy_t / dz_t = c_0$. To see the effect of z on y for 1, 2, … and j periods after the intervention, use recursive substitution in 6 to obtain y(t+1) as

$$y_{t+1} = a_0 + a_1 y_t + c_0 z_{t+1} + \varepsilon_{t+1} = a_0 + a_1(a_0 + a_1 y_{t-1} + c_0 z_t + \varepsilon_t) + c_0 z_{t+1} + \varepsilon_t$$
$$= a_0 + a_1(a_0 + a_1 y_{t-1} + c_0 z_t + \varepsilon_t) + c_0 z_t + \varepsilon_t \qquad (4.10)$$

since $z_{t+i} = z_t = 1$ for all $i > 0$. Hence, $dy_{t+1} / dz_t = c_0 + c_0 a_1$.

Clearly, the combined effect consists of the immediate effect of z(t+1) on y(t+1) given by c_0 and the effect of z(t) on y(t) also given by c0 multiplied by the effect of y(t) on y(t+1) given by a1. By successive substitutions in the same way, it is the j-step ahead impulse response of y(t) to z(t) is given by

$$dy_{t+j} / dz_t = c_0(1 + a_1 + ... + a_1^j). \qquad (4.11)$$

This is the short-run impact measure for different lags from the start of the intervention. Taking limits as j tends to infinity, and using the sum to infinity of the geometric series in the bracket $s_\infty = 1/(1-a_1)$, the long-run impact of the intervention becomes $c_0/(1-a_1)$. If 0 < a1 < 1, the impact of the intervention increases with j towards this long-run value, but if -1 < a1 < 0, the impact induces a damped oscillation of y(t) toward the long-run level $c_0 /(1-a_1)$.

The above notes used a single autoregressive time series of order 1 to explain intervention effects which typically combine effects of past and future values of y(t) and z(t) on the values of y(t) at different time lags. In a simultaneous equation VAR model, instead of one equation, each variable in the system is expressed as an autoregressive time series model of all the other variables, say of order p as follows.

4.3.4 Some notes on VAR analysis

Suppose it is not known which of two variables x and y are dependent or independent only that each variable potentially influences the other. Treating each variable symmetrically, we set up a 2-variable equation system as follows:

$$y_t = c_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + \varepsilon_{yt} \qquad (4.12)$$

$$y_t = c_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + \varepsilon_{zt} \qquad (4.13)$$

In this system, each variable is a function of its lag and that of the other variable and since the maximum lag is of order 1, it is called a VAR(1) model. The error terms are assumed to be white noises with means zero and respective variances σ_y^2 and σ_z^2. In matrix form, the equations can be rewritten as:

$$X_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix} \Leftrightarrow X_t = C_0 + A_1X_{t-1} + \varepsilon_t \qquad (4.14)$$

To extend the system to more than 2 variables *n* (for example the 6 variables modelled in this chapter) and more than one lags *k* so that each variable is a function of the past *k* lags of itself and the other variables, we add more terms in (4.10) to obtain

$$X_t = C_0 + A_1X_{t-1} + A_2X_{t-2} + A_3X_{t-3} + ... + A_kX_{t-k} + \varepsilon_t = C_0 + \sum_{l-1}^{k} A_lX_{t-l} + \varepsilon_t .$$

(4.15)

In this chapter, we use a 6-variable VAR(k) system which mimics equation 4.11 as follows:

$$\begin{pmatrix} RET_t \\ WOM_t \\ AGG_t \\ WB_t \\ ORAT_t \\ M_t \end{pmatrix} = \begin{pmatrix} c^{RET} \\ c^{WOM} \\ c^{AGG} \\ c^{WB} \\ c^{ORAT} \\ c^{M} \end{pmatrix} + \sum_{l=1}^{k} \begin{pmatrix} a_{11}^l & a_{12}^l & a_{13}^l & a_{14}^l & a_{15}^l & a_{16}^l \\ a_{21}^l & a_{22}^l & a_{23}^l & a_{24}^l & a_{25}^l & a_{26}^l \\ a_{31}^l & a_{32}^l & a_{33}^l & a_{34}^l & a_{35}^l & a_{36}^l \\ a_{41}^l & a_{42}^l & a_{43}^l & a_{44}^l & a_{45}^l & a_{46}^l \\ a_{51}^l & a_{52}^l & a_{53}^l & a_{54}^l & a_{55}^l & a_{56}^l \\ a_{61}^l & a_{62}^l & a_{63}^l & a_{64}^l & a_{65}^l & a_{66}^l \end{pmatrix} \begin{pmatrix} RET_{t-l} \\ WOM_{t-l} \\ AGG_{t-l} \\ WB_{t-l} \\ ORAT_{t-l} \\ M_{t-l} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^{RET} \\ \varepsilon_t^{WOM} \\ \varepsilon_t^{AGG} \\ \varepsilon_t^{WB} \\ \varepsilon_t^{ORAT} \\ \varepsilon_t^{M} \end{pmatrix} \quad (4.16)$$

where the superscript *l* identifies the matrix of interaction model coefficients A_l for each lag *k.* In this specification, the variables on the L.H.S. which are, respectively, retention, word of mouth, joined or did not join market aggregators, customer win-back strategies, other channels, and marketing spend, can potentially affect each other. The VAR model relates the values of these variables at time *t* to those at *k* time points in the past including the immediate time point *t-l.* The mediators of this relationship are a set of initial constants or average values of these variables captured as the first vector on the R.H.S. of the model, and a series of matrices with i-j elements that capture the different time-varying force-effects of one variable upon the other, and a vector of error terms which model the deviation of predicated and actual model results. Since these mediators are intrinsically *vectors* or matrices and the relationships are *regressions* of a bunch of variables on themselves and each other (*auto* regressions), the label vector autoregression (VAR) is applied to this species of structural time series models.

The VAR model assumes that the variables are stationary and integrated of order zero I(0), so that the system can be estimated using least squares applied to each question. However, if the variables are nonstationary, integrated of order one I(1), and not cointegrated in the sense that there is no relationship binding all of them, 'the variables are differenced once to make them stationary' (Hill, *et al.* 2012, p. 499-500). That is, the interrelations among them are examined using a VAR framework in their first differences.

If the variables in the model are non-stationary, I(1) integrated and jointly cointegrated in the sense that there is a linear relationship among them, then a special form of the VAR model is used. This special form of a VAR model introduces successive error term from the linear relationship and is known as a 'Vector Error Correction (VEC) model' (Hill *et* *al* 2012, p. 500-501). A general VEC model is of the form earlier specified in equation 4.5 above and recalled below (equation 4.17) for easy follow through,

$$\Delta Y_t = c + \sum_{i}^{k} \Phi_i \Delta Y_{t-i} + \Psi Z_t + \Lambda X_{t-1} + \varepsilon_t \tag{4.17}$$

In this study, it is not easy to know apriori the relationship among 6 marketing variables. Hence, a standard VAR model of the form in Equation 4.12 is used to estimate the impulse response function (IRF) effect of every variable on the others, which is the main focus of this research, particularly the effect of aggregators AGG(t) on other variables. Also, as argued in Sims (1980), VAR analysis should be conducted where necessary without differencing, since the main goal of the analysis is to determine the relationships among the variables, not the parameter estimates. This is because differencing ''throws away'' information about the possible comovements or cointegration relationships in the data sets.

Implementing VAR analysis involves a number of steps which are summarised in the empirical analysis below.

4.4 Empirical analysis and interpretations of modelling results

4.4.1 Data description and exploratory data analysis (EDA)

The data for this research was provided from an established UK car insurance company between the dates of Jan 2006 to August 2009. The data contained monthly figures of spend, quotes and sales from the different media channels of acquisition to be investigated. The different channels investigated are: Retention, Word of mouth, Aggregator, Win back, and Other direct channels (this includes all direct web and phone channel).

When gathering a quote from the car insurance company, the company requests the customer to choose where they had heard the company, which provides the Word-of-mouth and other direct channels acquisition routes; for aggregator, this information is fed directly via the price comparison site; retention uses the retained flag from the system, and for win back, the customer has to respond to an email sent to them. For customers that may have contacted the company on more than one different channel, e.g. directly and via a price comparison site, the data uses the initial contact channel, which also contains the initial contact date.

4.4.2 Data

When investigating the effect of aggregators on price comparison sites, certain customer acquisition channels need to be investigated: win-back quotes, where a customer had previously left and has since returned; word-of-mouth, when a customer had recommended the company; renewed, where the customer has renewed their car insurance; aggregator, where the customer contacted the company via an aggregator and finally, all the other direct channels to get a quote whether online or by telephone.

Table 4.1: Descriptive statistics for quote channels prior and post aggregator

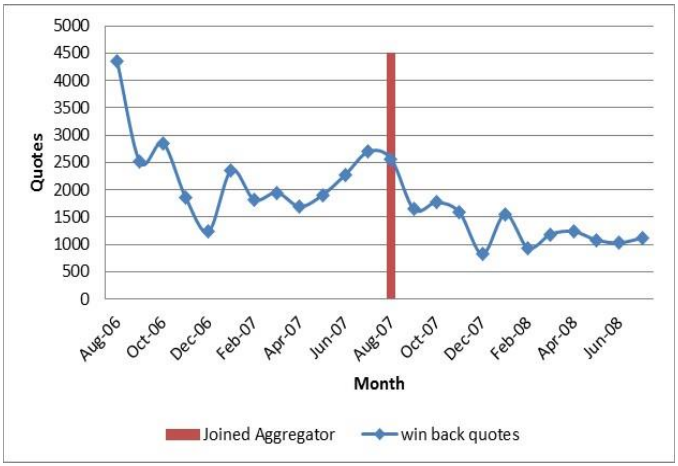
Month	Win Back	Word of Mouth	Other Direct Channels	Marketing Spend*	Average Premium	Renewal rate
Feb-07	1821	13453	161033	£857,210	£487	71.5%
Mar-07	1950	13933	169642	£970,420	£499	71.3%
Apr-07	1695	13738	149075	£518,864	£494	72.0%
May-07	1907	13999	143331	£527,128	£477	71.6%
Jun-07	2280	12854	123387	£415,724	£489	72.0%
Jul-07	2698	13523	122563	£295,158	£483	70.1%
Average	***2059***	***13583***	***144839***	***£597,417***	***£488***	***71.4%***
Joined Aggregator August 2007						
Feb-08	939	8316	61552	£317,871	£452	69.6%
Mar-08	1188	11841	79921	£293,482	£456	69.9%
Apr-08	1238	11042	75584	£285,194	£471	69.5%
May-08	1081	10786	73773	£264,978	£458	70.7%
Jun-08	1031	10481	68290	£170,769	£446	71.1%
Jul-08	1119	11146	67695	£297,372	£436	72.2%
Average	***1099***	***10602***	***71136***	***£271,611***	***£453***	***70.5%***
% Difference	**-46.6%**	**-21.9%**	**-50.9%**	**-54.5%**	**-7.2%**	**--1.3%**

*Marketing spend excludes any payments made to aggregators

Using a month-on-month comparison, seasonal effects can be disregarded. As can be seen from the above table 4.1, apart from renewal rates, aggregators have had a dramatic impact on all channels as well as marketing spend and the average car insurance premium. Win-back channel has dropped by 47%, word-of-mouth by 22%; other channels by 51%, marketing spend by 55% and average premium 7% (£30). This demonstrates that aggregators have a negative effect on all channels and can also affect car insurance premiums.

The above facts are demonstrated graphically below for a full two year period, from August 2006 to July 2008. The year prior to joining a price comparison site (2006) and the year following joining (2008) are used to demonstrate any changes recorded. It is noteworthy that aggregators had been established prior to the company joining aggregators and that the insurance company had also witnessed some degradation in their total quote volumes prior to joining.

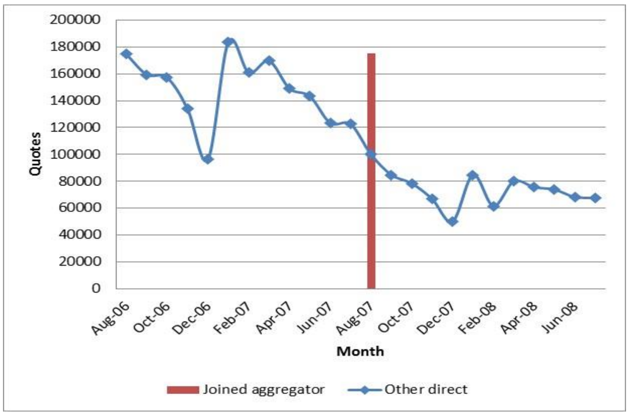
Figure 4.1: Descriptive Statistics of win-back quotes before and after joining aggregator



Observations	24
Mean	1836
Median	1733
Maximum	4355
Minimum	826
Std. Dev.	794
Skewness	1.40
Kurtosis	3.05

The company involved did not appear on aggregators when they first appeared so some degradation can be observed prior to joining. Since joining aggregators, it can be observed that the win-back channel is not as effective as it was.

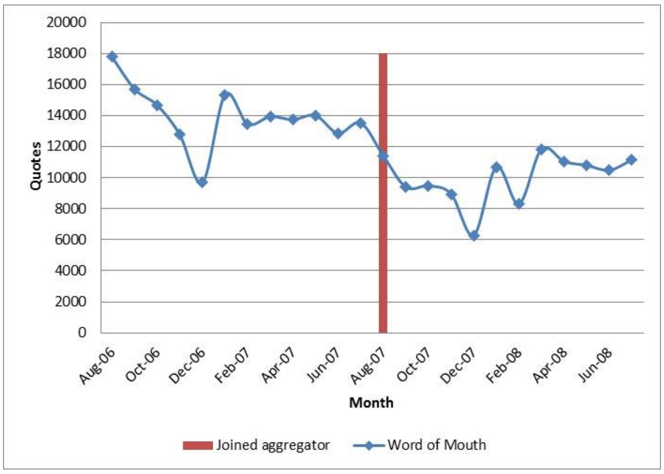
Figure 4.2: Descriptive statistics of other channel quotes before and after joining aggregator



Observations	24
Mean	111,071
Median	98,099
Maximum	183,757
Minimum	50,065
Std. Dev.	42362
Skewness	0.30

Prior to joining aggregators the 'other' direct channels were decreasing. After 6 months of joining, the trend had started to plateau. The graph above represents the main channel for contacting the company directly, thus the effects are more visible, especially when viewing quote volumes, Jan 2007 (183757) and Jan 2008 (84792) a decrease of nearly 100,000 quotes.

Figure 4.3: Descriptive statistics of word of mouth quotes before and after joining aggregator



Observations	24
Mean	11,968
Median	11,626
Maximum	6,287
Minimum	17,777
Std. Dev.	2,665
Skewness	0.06
Kurtosis	-0.09

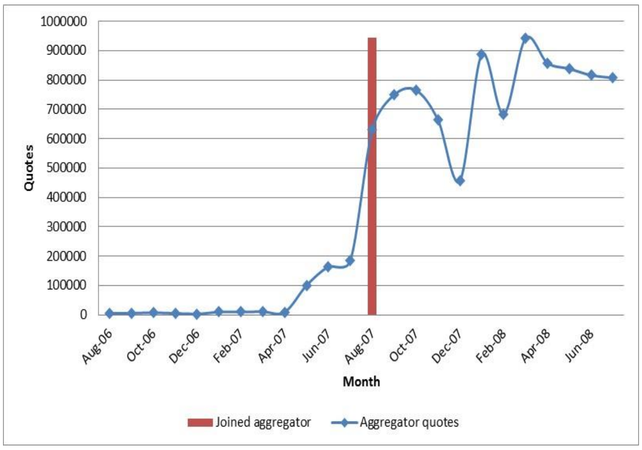
The word of mouth metric has been one of the least affected in terms of in the number of quotes since joining aggregators. This could be due to the rise of social media, such as Facebook, that encourages word-of-mouth recommendations.

Figure 4.4: Descriptive statistics of renewal rates before and after joining aggregators

Observations	24
Mean	71.38%
Median	71.13%
Maximum	75.06%
Minimum	69.50%
Std. Dev.	0.01
Skewness	0.78
Kurtosis	1.11

From the graph above a slight decrease in renewal rates can be observed. If customer think that their renewal price is cheap enough, then this may not cause them to investigate other companies and more efficiently with lower search costs, using aggregators.

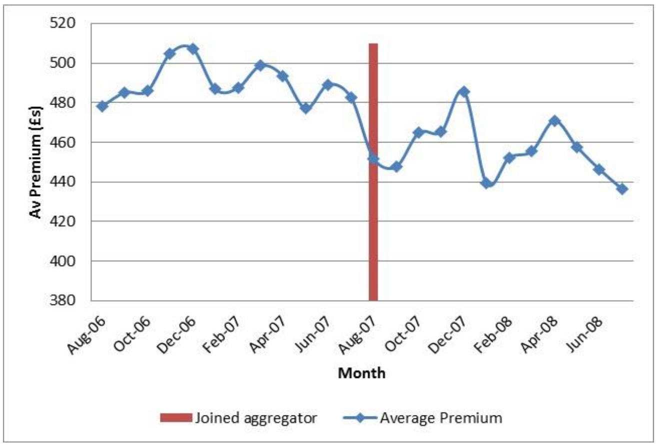
Figure 4.5: Descriptive statistics of aggregator quotes before and after joining aggregators



Observations	24
Mean	400,633
Median	320,817
Maximum	941,951
Minimum	3,189
Std. Dev.	379,616
Skewness	0.12
Kurtosis	-1.93

Prior to joining the aggregator, the company was testing its quote process to establish whether its systems could manage with the increase in quote numbers. August 2007 is when the company went fully live with aggregators. Before this, the main acquisition channel (other) had its largest number of quotes (183,757) in January 2007 (largest number). However, aggregators managed to gather 941,951 quotes in March 2008. This clarifies why the company had to test its quote engine process and why the other acquisition channels have been affected so dramatically.

Figure 4.6: Descriptive statistics of average premium before and after joining aggregator



Observations	24
Mean	£472
Median	£477
Maximum	£507
Minimum	£436
Std. Dev.	20.80
Skewness	-0.18
Kurtosis	-1.07

As well as quotes, average premium can also be seen to be affected. The insurance company had not joined aggregators as soon as they were formed and decided to wait. The effect of aggregators can be seen at the beginning of the time period chosen in the graph, with a slight downward trend. When the insurance company decided to join aggregators, did their average premiums truly started to decrease. The decrease in premium could be attributed to the way that customers can easily compare different companies, where being the 'cheapest' means a better placing on the aggregator site.

4.4.2 VAR Test results

The reason for the numerous tests required before producing the IRF is that if there exists unit roots and/or cointegration, then the IRF loses its predictive power.

Step 1

The first part of the model developments used the AIC and BIC tests to verify the lag order and to also check that a VAR model would produce the necessary results, instead of, for example, a vector auto regressive moving average (VARMA) model. The results verified presented in Tables 4.2 and 4.3 show that a lag order of 4 should be used.

Table 4.2: AIC results for the lag order of the model

Order 1	Order 2	Order 3	Order 4
AIC -20.2714	AIC -20.5659	AIC -20.6251	**AIC -21.2219**

Table 4.3: BIC Results for the order of the model

Lag	**MA 0**	**MA 1**	**MA 2**	**MA 3**	**MA 4**	**MA 5**
AR 0	13.85663	13.82101	13.76469	13.83584	14.61243	13.76066
AR 1	19.67225	19.16311	18.59353	18.45116	18.85895	17.42494
AR 2	19.88681	19.3296	18.60412	17.97094	17.91678	16.01783
AR 3	20.06206	19.24973	18.17964	16.82996	16.49364	13.22352
AR 4	**20.71461**	19.45398	18.0062	16.36999	13.38679	9.198823
AR 5	19.62387	17.97812	16.02435	13.10496	8.796386	1.834561

As shown in table 4.2 and table 4.3 the tests show that the model should have a lag order 4 due to order 4 having the lowest value. Using the BIC results, table 4.3, it can also be shown that a VARMA model is not required (due to the highest value being found at AR4 MA0) and a VAR model should be the one to use.

Step 2

The second part of the analysis tested for stationarity for each endogenous variable. The Dickey-Fuller test along with the KPSS found that they are not stationary and seasonal differences will need to be taken. In Addition, the KPSS test statistic provides proof that we cannot rely on the Dickey-Fuller Unit Root test alone. The tests are presented in Tables 4.4, 4.5 and 4.6 below.

Table 4.4: Dickey-Fuller unit root test

Variable	**Type**	**Tau**	**Pr < tau**
Other	Zero Mean	0.89	0.8981
Other	Single Mean	0.49	0.9853
Other	Trend	-1.10	0.9211
Word of Mouth	Zero Mean	-0.63	0.4403
Word of Mouth	Single Mean	-1.45	0.5537
Word of Mouth	Trend	-1.08	0.9249
Win back	Zero Mean	0.26	0.7580
Win back	Single Mean	-2.70	0.0797
Win back	Trend	-2.73	0.2280
Aggregator	Zero Mean	0.48	0.8154
Aggregator	Single Mean	-1.43	0.5626
Aggregator	Trend	-2.70	0.2398
Retention	Zero Mean	-0.65	0.4307
Retention	Single Mean	-2.84	0.0580
Retention	Trend	-3.50	0.0470
Spend	Zero Mean	-1.29	0.1791
Spend	Single Mean	-0.31	0.9169
Spend	Trend	-2.79	0.2059

The results from Table 4.4 shows that the retention channel may be stationary, the KPSS test will be carried out to verify this. The null hypothesis of the KPSS states that the time series is stationary. As the model has an intercept, it performs two tests: single mean (mu) and trend (tau). The null hypothesis of stationarity is rejected if the KPSS test statistic exceeds the respective critical value.

Table 4.5: KPSS test

Variable	ETA (mu)	ETA (Tau)
Other	0.8131	0.4231
Win back	0.3519	0.2099
Word of Mouth	0.6407	0.3963
Aggregator	1.1765	0.1240
Retention	1.1365	0.1391
Spend	1.5539	0.2399

Table 4.6: KPSS critical values

Type	Prob10pr	Prob5pr	Prob1pr
Mu	0.3470	0.4630	0.7390
Tau	0.1190	0.1460	0.2160

From tables 4.5 and 4.6 the stationary and linear trend all exceed 10%, so we can say that they are not stationary and seasonal differences will need to be taken. The KPSS test statistic provides proof that we cannot rely on the Dickey-Fuller Unit Root test alone.

Step 3

The third section involved testing for the cointegration. If cointegration exists then error correction terms are added to correct the cointegrating variables depending on its trend. The tests results had shown that there is a separate drift and no separate linear trend. Due to evidence of cointegration, a Vector Error Correction model (VECM) will need to be developed as in equation 4.5.

To test for any exogeneity between the endogenous variables, after the VECM was created the Granger-Causality Wald Test was carried out. By rejecting the null at 10% mark, we conclude that there are no significant exogeneity effects upon each of the variables, so the impulse response function can be developed from the parameter effects. The test results are presented below.

Cointegration test results

If cointegration exists, then error correction terms are added to correct the cointegrating variables depending on its trend. The following alternative test scenarios are explored:

- Case 1 - there is no separate drift
- Case 2 - no separate drift, but a constant enters with the error correction term
- Case 3 - a separate drift with no separate linear trend
- Case 4 - a separate drift with no separate linear trend, but a linear trend appears with the error correction term
- Case 5 - a separate linear trend

(SAS, 2008)

Tables 4.7 and 4.8 demonstrate that there is cointegration, either Case 2 (the hypothesis H0) or Case 3 (the hypothesis H1) and the significance levels need to be considered (table 4.8).

Table 4.7: Cointegration rank test using trace

H0: Rank=r	**H1: Rank>r**	**Eigenvalue**	**Trace**	**5% Critical Value**	**Drift in ECM**	**Drift in process**
0	0	0.6179	152.706	93.92	Constant	Linear
1	1	0.5053	91.1282	68.68		
2	2	0.2628	46.0871	47.21		
3	3	0.2427	26.5724	29.38		
4	4	0.1281	8.7779	15.34		
5	5	0.0001	0.0065	3.84		

Table 4.8: Cointegration rank test using trace under restriction

H0: Rank=r	**H1: Rank>r**	**Eigenvalue**	**Trace**	**5% Critical Value**	**Drift in ECM**	**Drift in process**
0	0	0.6189	154.9596	101.84	Constant	Constant
1	1	0.5057	93.2272	75.74		
2	2	0.2684	48.1334	53.42		
3	3	0.2446	28.1301	34.8		
4	4	0.1281	10.176	19.99		
5	5	0.0217	1.4026	9.13		

From table 4.8 the cointegration rank is chosen to be 1 by the result previously, and the p-value is 0.8353, the Case 2 can be rejected at the significance level of 10%. So we can say that there is a separate drift and no separate linear trend (case 3). Due to evidence of cointegration, a Vector Error Correction model (VECM) could be developed, but as argued in the notes on VAR and VEC modelling, for interest in determining relationships among the marketing variables, a standard VAR model is adequate.

Table 4.9: Hypothesis test of the restriction

Rank	**Eigenvalue**	**Restricted Eigenvalue**	**DF**	**Chi-Square**	**Pr > ChiSq**
0	0.6179	0.6189	6	2.25	0.895
1	0.5053	0.5057	5	2.1	0.8353
2	0.2628	0.2684	4	2.05	0.7272
3	0.2427	0.2446	3	1.56	0.669
4	0.1281	0.1281	2	1.4	0.4971
5	0.0001	0.0217	1	1.4	0.2374

Table 4.10: Granger-Causality Wald test

Test	**DF**	**Chi-Square**	**Pr > ChiSq**
1	20	58.63	<.0001
2	20	59.02	<.0001
3	20	95.73	<.0001
4	20	72.71	<.0001
5	20	140.21	<.0001
6	20	31.21	0.0525

The complete VAR model is displayed in appendix 4.2.

4.5 Empirical Results

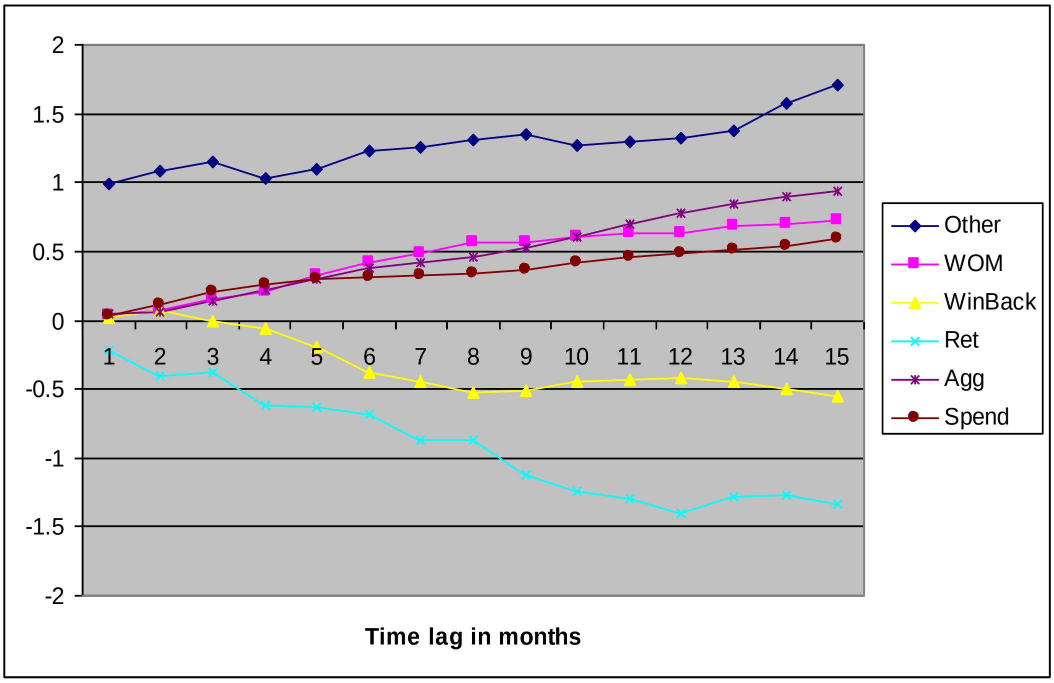
4.5.1 Using impulse response functions to measure the impact of price comparison sites on marketing mix components

To measure the impact of the different customer channels (endogenous variables) on each other, impulse response functions (IRFs) are used. IRF's can be used to graphically represent a time path of the effects of dependent variables on other variables. The IRF will be used to investigate the effect of one of the variables towards another when an unexpected shock enters the system (introduction of price comparison sites). Adopting a 15-month lag, the short and long term effects can be measured.

The graphs will use an accumulated IRF, so we can monitor the trend over a particular period of time for each of the different channels. Should the shock become stable then we would expect the graphs to level off and plateau, any increase or decrease in the gradient of the graph would demonstrate a longer lasting effect. The elasticity between the variables is comparable, as they have been transformed with a logit function.

Short and long-term effects

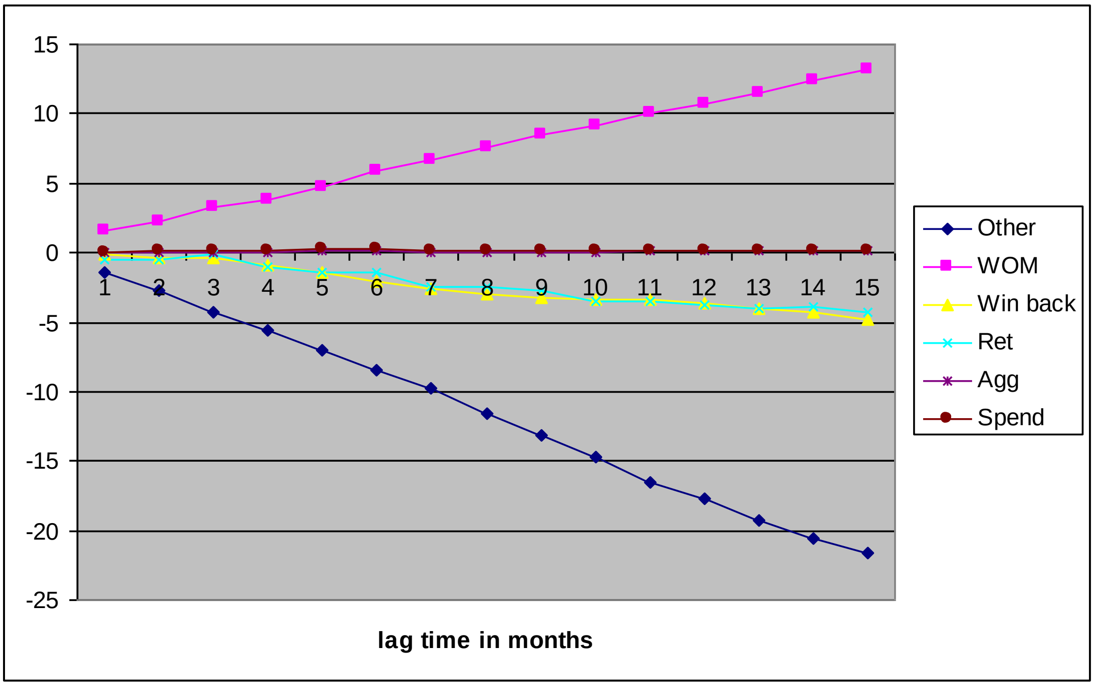
Figure 4.7: IRFs for other channel ratio effects



The other ratio effects can be considered as the main direct channels for the customers. From Fig. 4.7 it can be seen that customers acquired via other direct routes, have a long lasting positive effect on word-of-mouth, aggregator, spend and on itself, due to the gradient of the graph increasing. The spend increase confirms results in Ambler (1997, pp.290) that ‘demonstrating effective leads not only present budgets being better spent but bigger budgets being made available’. Fig. 4.7 also provides insight into how new acquisition targets behave differently with respect to retention and win-back. As for both retention and win-back, both types of customers have heard of the company before, the figure demonstrates different attitudes to new and old customers. This demonstrates that customers who regularly shop around for their car insurance may influence others for new acquisition, rather than going back to their old company or renewing their car insurance.

The majority of the effects can be considered weak, when considering the scale of the axis.

Figure 4.8: IRFs for Word of Mouth (WOM) advertising

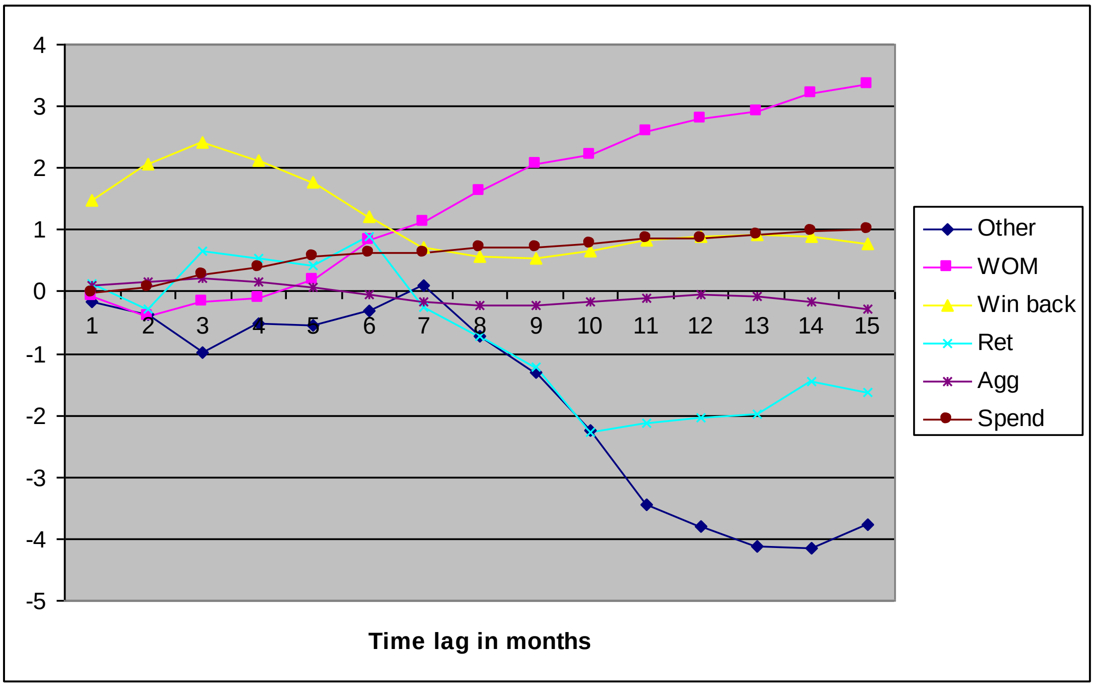


From the Fig. 4.8, the customers acquired through WOM have a strong positive effect on future word-of-mouth acquisition. The other channels react negatively, with the other channel exceptionally negative, which differs from the view in Trusov *et al*. (2009) that there are increases in acquisition across all channels when using word-of-mouth. As expected, word-of-mouth acquisition has virtually no effect on marketing spend.

WOM also has little or no effect on aggregator acquisition rates, which demonstrates that people who contact the company via word-of-mouth tend to contact the company directly rather than compare the company on aggregators.

WOM provides a good value channel for businesses, and its strong acquisition rate takes the good leads from the other channels, thus having a negative impact on these channels

Figure 4.9: IRFs for Win Back



Win Back customer are communicated with via email campaigns. The customers are targeted between 11 and 12 months since leaving the company so the timings of the campaign should target the customer when they are in the market of acquiring car insurance.

From Fig. 4.9, the effects of Win Back acquisition seem quite erratic short term within the first 6 months, but the long-term effects demonstrate a strong negative effect on the other channel acquisition rates, a negative effect on retention rates and a slight negative effect on aggregator rates. What is noted is the strong positive effect on the WOM channel, which demonstrates that customers that have been won-back may influence their friends to purchase their car insurance from the same company.

Gee *et al* (2008, p.370) note that ‘A win-back strategy is recommended as previous customers are less costly to win-back compared to the costs of acquiring of new customers’. We can also extend this to include that they also influence word-of-mouth acquisitions, which are free, and may stop customers from using price comparison sites which saves even more money for the company.

Figure 4.10: IRFs for Retention channel

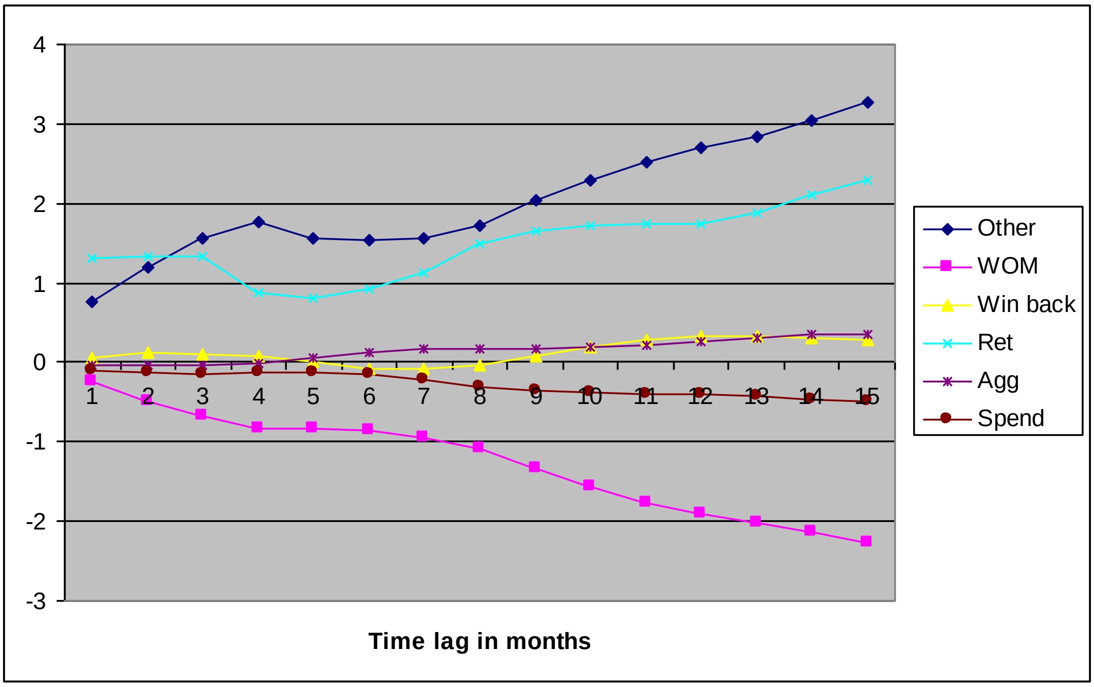
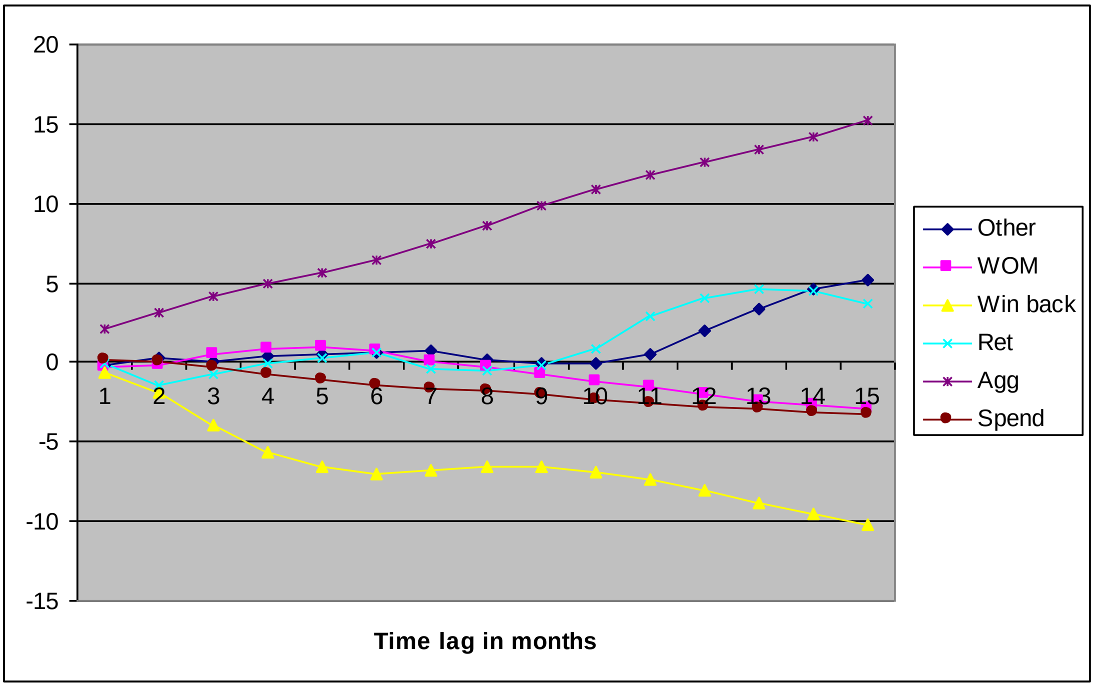


Figure 4.10 demonstrates that retention has a slight positive effect on aggregator and win-back rates, which shows that people who stay with the company may slightly influence those people who purchased via price comparison sites and returning customers.

Retention also has a strong positive effect on other customers who stay with the company and renew their insurance. This may show that people who stay with the company tend to influence the other direct channels and those customers that renew, tend to keep renewing. For customers who renew, they tend to stay away from price comparison sites compared to people who use the other direct channels. Customers who renew have a negative effect on word of mouth acquisition rates, which may mean that even though the customer stays with the company, they may not be telling other people.

The graph also agrees with the common mantra that it is cheaper to retain a customer than acquire a new customers (Rosenberg and Czepiel, 1984; Blattberg and Deighton, 1996) as the effect marketing spend is negative.

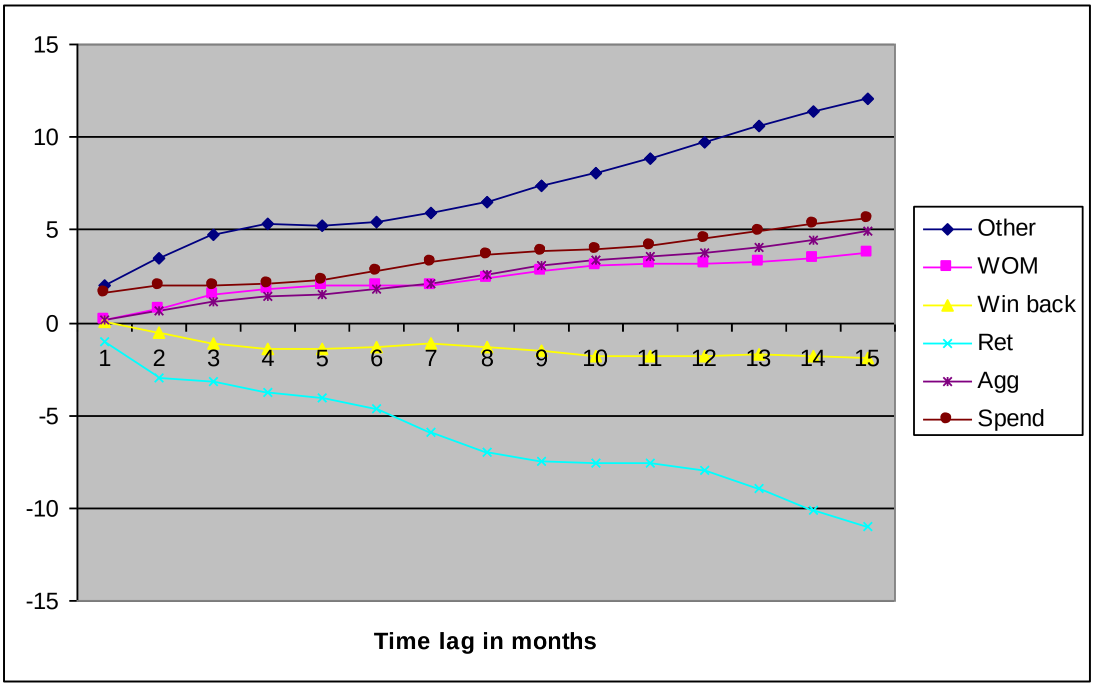
Figure 4.11: IRFs for Aggregator channel



As it is the price comparison sites marketing that brings the customer to their site and not to the company directly, 'the indirect measure of [brand] equity… should be linked to objective measures of consumer choice whenever possible' (Leuthesser *et al.*, 1995 pp.65). Figure 4.11 shows the effect of customers who use aggregator sites and are given customer choice. Figure 4.11 demonstrates that customers acquired via price comparison sites tend to have a strong positive over other users of aggregators sites, which also demonstrates strong brand equity. After 12 months, price comparison sites have a positive effect on retention and other ratios, which shows that some customers may not use the price comparison site when their renewal for car insurance appears.

The cost of customers purchasing via aggregators comes from the marketing budget, so the negative impact is demonstrated, i.e. as more people are acquired through the price comparison sites, this cost is taken from the marketing budget. The graph also shows that certain win-back customers would prefer to try out a price comparison site rather than going straight back to their old insurance company.

Figure 4.12: IRFs for Marketing Spend



The results show that the marketing spend seems to be concentrated mainly on acquisition channels rather than retaining and winning back lost customers. The strong effect on acquisition is clearly seen due to the continuing increase in the gradient of the slope. The graph also shows that its strategy may be based on building customer size based on acquiring customers from other companies.

The results differ from Yoo, S., and Hanssens D. (2005) who found that retention rates are not affected by advertising, but as we can see the marketing spend (the advertising) has a strong negative affect on customer retention.

Figure 4.12 also mimics the behaviour as in Figure 4.7 with new acquisitions always on the increase and retention/win-back on the decrease. This demonstrates a strong link between advertising and direct channel customer contact, which is expected.

4.6. Summary and conclusion

4.6.1 Discussion of the results in light of the research objectives

The main purpose of this chapter and study is to increase the understanding of the effect of aggregators within the UK car insurance industry. This research results presented above provides an insight of how re-intermediation relate to strategic marketing planning and implementation via the marketing mix, in helping us to reposition the case company with regards to its future growth and profitability. This objective will be explored in more detail in Chapter 6 of the thesis which will focus on business modelling. The following points relate the findings to previous work and present further insights which will inform the in subsequent analysis chapters, for example Chapter 7.

Firstly, the market response model findings agree with Pauwels and Neslin (2008) that adding a new channel does affect customer retention. Aggregators make it easier for the consumer to shop around and get the best deals when considering car insurance, but the results provide insight that customers who used aggregators initially, may be more inclined to renew.

Aggregators have a strong effect on the main acquisition channel. This may demonstrate that customer may be going to the company directly first to get a quote, then use the price comparison site to compare prices. People who shop on-line are more impulsive and are more responsive to direct marketing (Brashear *et al*., 2009). If the insurance company used their marketing spend on DM, this may limit the number of customers using aggregators, thus saving the company the commission it has to pay for each sale to the price comparison site.

The market response models demonstrate the cannibalising effect on win-back and word-of-mouth with the introduction of price comparison sites, which are both cheap acquisition channels. These channels are not the main sources of customer acquisition, but do demonstrate that price comparison sites do not affect the car insurance sites positively as a whole. It is also worth noting that aggregators have a negative short-term, but a positive long-term effect on customer retention. This may be because a customer who already purchases a price-competitive quote through an aggregator will be more likely consider the insurance company cheaper than competitors in the future, having observed the differences in prices at the time of searching for the competitive quote.

This research shows that aggregators negatively impact win-back customers, but have a positive effect on retention and support other channels' acquisition, although these effects are only apparent in the long term. Also, this research shows that win-back, retention and word-of-mouth require little spend so represent a good source of customers. Finally, the results also show that traditional marketing is still needed to generate new customers.

Car insurance companies in the UK should not be treated as a standard company. Due to the legal requirement needed to purchase car insurance, the companies must be aware of their ideal customers. Aggregators can give the car insurance companies access to a wide range of customers if they choose to adopt them. Thus, this work may contain some important implications for such companies considering the usage of price comparison sites. It may provide insights into the effect of aggregators on customer acquisition and the effects of marketing spend on retention and other channels.

4.6.2 Future research and limitations

This chapter has some limitations for future work. Firstly, this chapter only considers those companies that have joined an aggregator, so neglects the use of such customers who chose not to join aggregators. Secondly, this chapter only considers one car insurance company; it does not review the impacts of aggregators for other car insurance companies. Lastly, this research only considers the car insurance market and neglects other insurance products that may be affected by aggregators.

Chapter 5: Price comparison sites, car insurance business modelling statistical analysis and scenario modelling

5.1. Introduction

Financial price comparison sites (aggregators) are instigating considerable changes within the UK car insurance industry. Regardless of whether a car insurance company decides to join an aggregator, or not, their impact has changed how their business operates. To fully understand the impact of aggregators within the car insurance environment, business models can be used. Business models break down complicated components of the business into easier to understand processes. Even though business models may lack 'theoretical grounding in economics or in business studies (Teece, 2010, p.175), used correctly they can be used to look at the business processes for creating value (Petrovic *et al*., 2001). This chapter provides analytical support for this logic and economic value with regards to the case company.

Although business models have been researched for different scenarios, there is little research on the effect of the adoption of an aggregator within the UK car insurance industry. The claims processes for car insurance have been researched before (Telang & Singh, 2009; Oliveira *et al*., 2007), but aggregators do not directly affect these processes, so cannot be further considered. This leads to an absence of a business model detailing the effects of aggregators, which needs to cover the aspects that are directly affected by aggregators. In modelling these effects in this study, the implications within the business model will be observed, such as sales, customer retention, return on investment and marketing.

Investigating different aspects of the business model such as marketing, sales, retention and return on investment, the effects were monitored by using different statistical modelling approaches. An application within the case company reveals that aggregators have a dramatic effect on the number of customers contacting the company as well as its return on investment (ROI). Hence, the chapter uses general linear models as well as time series techniques such as autoregressive, moving averages, autoregressive integrated moving average and GARCH models, in order to explore the effects of insurance companies implementing aggregators. The results have strategic implications for the UK car insurance industry, with regards to whether or not companies should use aggregators in their distribution channel mix, by using key aspects of the new business model.

The main focus of this chapter is to begin to develop a business model which includes the implementation of aggregators. This will be achieved by reviewing relevant literature regarding business models within the insurance industry and other industries that are affected by aggregators. The second aim is to map different scenarios with regards to the adoption of a price comparison site, and the effect this will have upon the business model. The final aim is to provide further insight into the sales, retention and return of investment by statistical modelling techniques.

The outcomes of the research will produce a deeper understanding of the case company and UK car insurance that will benefit future research, senior managers within the car insurance industry and other industries considering adopting an aggregator to their business.

The remainder of this chapter has the following layout. Section 5.2 reviews the general literature for the development of the business model which will cover different types of business and will also need to include the role of aggregators. Section 5.3 discusses the theoretical background to the statistical models used in providing the necessary insight into the business model. Section 5.4 presents the empirical analysis and the interpretation main results and concludes the chapter.

5.2. A business model framework and alternative scenario development

Research into companies adopting new technology and encompassing a new channel into their marketing mix has been accomplished using a variety of theoretical frameworks. Reviewing relevant different business models based on relevant research will help construct a business model (Figure 5.1). From this review, the key constructs of the business model can be defined, using the relevant research to support the outcomes.

Aggregators are still a relatively new phenomenon and their emergence within the UK car insurance has yet to be fully researched. Firstly, by reviewing the price comparison site industry, this chapter will be able to delve into the effect they have had within the UK car insurance industry.

5.2.1 Price comparison sites

Since the advent of the internet, the acquisition process of most products has been affected, with aggregators changing the purchasing landscape even more dramatically. Aggregators have evolved from screen scraping 'whereby the aggregator accesses the target site by logging in as the customer, electronically reads and copies selected information from the displayed webpage(s), then redisplays the information on the aggregator's site' (E-banking, nodate), to using XML which connects straight to the car insurance quote engines. This produces a faster response for the customer, thus saving the customer even more time. The rise of the price comparison sites being the channel of choice of the customers is seen in the fact that they 'instigated 25% of all private motor insurance sold in 2007' (BBC, 2008) and in 2010, 73% of online consumers had used a price comparison site (OFT 2010).

Aggregators allow the customer to view many prices from many different car insurance providers at the same time. This process reduces the customer's buyer search costs. This transparency of costs can benefit the customer in three ways:

- Search costs decrease as more information is made available at no additional cost
- The value of a purchase is more apparent
- Information may become available that allows a consumer to transact at a lower price for a given product.

(Granados *et al*., 2006, pp.154-155)

The decrease in search 'costs' can also be seen in the time saved which would have been spent searching numerous companies either by phone or the internet. The value of a product increases, as a product may offer more benefits for the same amount of money, this being the decisive factor for a potential customer. The transparency of the aggregators may also encourage companies to lower their prices, both to attract new customers and to retain current customers.

Aggregators make their money when the customer purchaces their insurance via the comparison website and the insurance company gives the aggregators a certain rate. The rate charged differs between the different insurers' companies (Simon, 2011). Usually, the insurance companies have to sign a contract saying that the price viewed on the price comparison site, is the same as if the customer went directly to the insurance company themselves. Competition between the aggregators is prevalent, each site wanting a particular company to be on their site and on no one else's. Due to this scenario customers would have to use more than one price comparison site if they wish to probe more of the car insurance market.

5.2.2 Business models

The ability to break down complicated components into easy to understand processes is the main strength of business models. Different industries require different types of business models, as there is no ‘one size fits all’ scenario. For this reason, a selection of papers was used to develop the proposed IBMR business model in this research: (Osterwalder and Pigneuir, 2002; Shafter *et al*., 2005; Bouwman *et al*., 2005; Chesbrough and Rosenbloom, 2002; Haaker *et al*., 2006; and Hamel, 2002).

Although the research papers listed above use different components, they all tend to agree on the importance of creating value for the brand and for the customer. They also provide valuable insight into the development of a business model. This research will use a four-segmentation approach (Figure 5.1) with an emphasis on a car insurance company joining an aggregator: value proposition; value relationship; customer relationship, and financial costs. The four main segments will contain module areas that make up the main segments.

Figure 5.1: Business model components

Value Proposition Product/Price Marketing the product/brand Distribution channels	**Value Relationship** Value chain Relationship building with partners Infrastructure
Customer Relationship Locating profitable customers Creating relationship with customers Satisfying customer needs	**Financial Costs** Costs of the other segments Profitability Competitor cost strategy

The business model needs to be adaptable to sudden shocks in the market. The business model proposed in this research will provide a framework for the UK car insurance industry within a price comparison site environment, and using the case company as the focus of related statistical analysis. It is these four segments that will provide the basis for the statistical models, which will be discussed in more detail in section 5.3.

The next section discusses the key segments within the business model: value proposition; value relationship; customer relationship, and financial costs. These four segments also provides the basis for the scenarios proposed within the following section.

5.2.3 Value proposition

In the UK, car insurance is a legal requirement under the Road Traffic Act 1988. The more comprehensive the insurance cover, the greater the premium. Within these levels, there could be a monetary excess to be paid by the customer towards any claims costs. The greater the excess the lower the customer should expect to pay for their car insurance and vice-versa. Within a price comparison site scenario, the different levels of cover and excess can make a car insurance company appear at the top of a price comparison site with the cheapest price, but with fewer benefits. The pricing strategy within car insurance can affect the type of customer the company attracts. Comparing price sensitive customers with loyal customers, 'the former can expect to pay a lower price while, somewhat counter-intuitively, the latter can expect to pay a higher price' (Morgan *et al*., 2006, p.135). So a company could increase their rates without losing loyal customers, but they should expect the price sensitive customer to leave or negotiate any further premiums.

Customers prefer to get value for their money, and although initial low prices may first attract the customer, it is the further benefits which are more important (Granados *et al*., 2008). This scenario does not however mean that companies can charge as much as they want, as customers will 'balance the benefits of the purchase against the costs' (Grewal *et al*., 1998, p.56). If a company chooses to target the price sensitive customers only, this would not suit customers who need a more comprehensive cover. This would create a dilemma for the car insurance company as price sensitive customers may be ideal for quick growth, but for long profitable relationships then a more expensive but comprehensive product should be considered.

The way a company targets their customers can provide an insight into how they conduct their business. 'The creation of a customer through marketing and innovation that manages a business must always be entrepreneurial in character' (Drucker, 2007, p.40). A company must therefore consider the pros and cons when considering whether or not to join a price comparison site and the effect this will have on their marketing strategy.

Marketing gives the company its identity, its equity. The brand equity gives the company its value e.g. 'when certain outcomes result from the marketing of a product or service because of its brand name that would not occur if the same product or service did not have that name' (Keller, 1993 pp.1). People's perception of a company is mainly derived from marketing, so the more familiar costumers are with the company, through advertising, the more likely they will trust the brand (Keller and Lehman, 2006).

If a company chooses to join a price comparison site, then their marketing strategy will need to change as they could be proved incorrect in quoting that they are 'the cheapest' quite quickly. As aggregators tend to sort the insurance companies by price, how a company propositions itself, its brand equity, will become important. When a company has a strong brand it becomes less vulnerable 'to competitive marketing actions, …[and] greater intermediary co-operation' (Delgado-Ballester and Munuera-Aleman, 2005, p. 187). A company with strong brand equity could gain customers from companies with a weaker brand equity (Leuthesser *et al*., 1995; Lim *et al*., 2012), so the marketing strategy will need a clear set of goals 'derived from the corporate strategy' (McDonald *et al*., 2001, p.342). Adjusting the marketing to a branding perspective from a direct perspective puts the company in a more competitive, stronger position from both a price comparison site and direct-approach point of view. This section leads to consider the scenario:

Scenario 1: - the effect on marketing if a company does not join a price comparison site

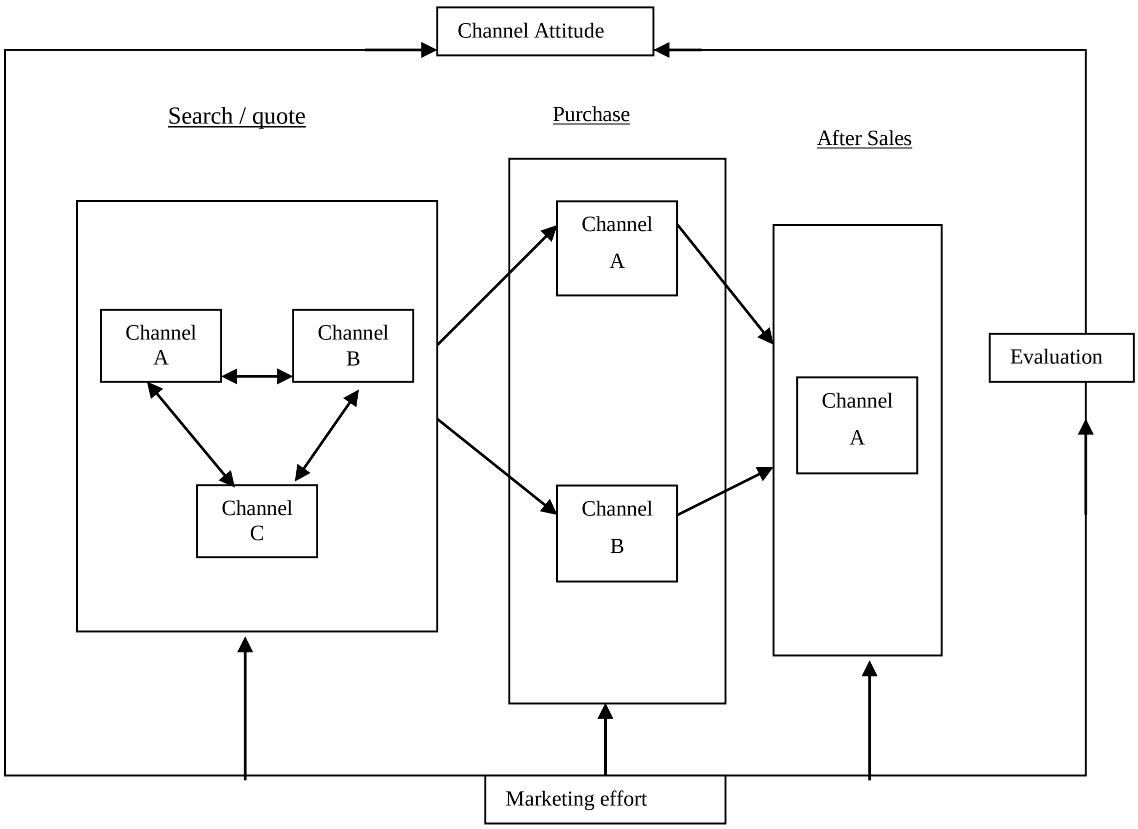
Marketing within the insurance industry provides the customer valuable information about the company itself, the service they provide and how to contact them. This scenario will provide insightful information in how the insurance company should position itself in the future.

5.2.4 Value relationship

The value relationship segment considers the process of creating relationships with the aggregators and how aggregators can affect the customer relationship. The aggregator would affect the direct relationship a customer may have had previously when contacting the insurance company to get their car insurance, by adding an extra link of chain between the customer and the insurance company.

If a new channel is introduced into the mix, the company will need to consider how people search, purchase and behave after sales. The researcher adapts the general model of customer choice (Blattberg, *et al*., 2008) to include a triple channel strategy instead of a duo channel strategy as follows.

Figure 5.2: Framework of triple acquisition channel strategy



Key: Channel A = telephone, Channel B=own web site, Channel C = price comparison site

Figure 5.2 demonstrates how a consumer can use multiple channels for their research to get a car insurance quote. A customer can easily phone, visit the company's website, go to a price comparison site or use a combination of all three to get a quote for their car insurance, but can only purchase from one of the channels of the insurance company. The way that the customer can only purchase from the insurance company can be likened to a car sales person showing the car, taking the car for a test drive, then before the contract is signed, the car sales leaves the customer to complete the process themselves. This leaves the aggregator in a strange position as being able to complete the sales process themselves. For the case company, after sales is conducted by telephone only, so for a customer to make any amendments to their policy or if they need to report a claim, this is conducted by phone, as shown in figure 5.2, after sales.

To drive people to make contact with the insurance initially, we would need to consider the marketing. Without the insurance company marketing, the customer would not have contacted them. Marketing create the opportunities for the customers to become aware of the company and its purpose. Marketing also affects the search/quote, purchase and after sales segments of Figure 5.2. To fully understand the marketing impact, customers are asked 'where did you hear about the company?' This leads to the '**evaluation**', which provides insight for the marketing department to enable their marketing strategy to be more profitable.

The final part of the diagram specifics the channel attitude. This segment encompasses the marketing activity and the distribution channel of choice. This will allow company to budget their staffing levels more accurately, that there are enough staff to maintain the website, as well as to answer the important telephone calls.

Cooperation between a car insurance company and an aggregator is an important tool when building a relationship. Using a cooperative approach benefits both parties involved, more than working independently (Bruhn *et al*., 2013), or as the common saying goes 'a problem shared is a problem halved'. If either the price comparison site or the car insurance company thinks that they could function better alone, then this may cause some conflict. As well as cooperation, trust is also required. In some instances a car insurance company may need to provide the price comparison site with the full figures of purchases made via their website, so that billing can be self-regulatory. In this scenario, 'trust becomes even more important for relationship and loyalty development' (Lages *et al*., 2008, p.688).

When a company joins a price comparison site they need to consider the organisational arrangements. There are two such arrangements:

- the closed model, in which there is a relatively fixed consortium of partners,
- the walled garden model, in which new partners are only allowed to join the value network if they comply to a certain rule

(Reuver and Haaker, 2009, pp.242)

The main reason why a UK car insurance company may wish to join an aggregator is that they may give the insurance company wider exposure in the market. The most common jurisdiction for a price comparison site is that the quote you get from them would be the same as if you contacted the car insurance directly. If this rule does not satisfy the car insurance company then this may cause some conflict between both companies. A price comparison site enforcing a walled garden approach, may give a sense of exclusivity, but since they make money by customers purchasing via their web site, this could mean losing money and customers using other sites to get a fuller range of quotes from more UK car insurance companies.

It is not in the best interests for an aggregator to favour one insurance company over another, as this may drive other companies away from the website; therefore the price comparison site needs to be fully transparent. A fully transparent site will need to provide accurate information about the product, as well as price (Granados *et al*., 2006). Having full transparency will strengthen the comparison site by attracting more customers, however one adverse effect could be that the company may lose sales due to the customer being more informed (Porter, 2001).

The relationship between the aggregator and the car insurance company must involve many different departments to see how the two companies can be aligned. Communication between all parties involved in setting up the compatibility between a price comparison site and a car insurance company is very important. As more information is shared between the members of a supply chain, this reduces the uncertainty and enhances the performance between the suppliers (Sirinivasan *et al*. 1994). The flipside to this argument would be that if either party withholds information from the other, this could lead to delays and a poor working relationship.

On the basis of the previous section we would need to consider the following alternative scenario, with a recall of the first scenario for easy follow through:

Scenario 1: - the effect on marketing if a company does not join a price comparison site

Scenario 2: - the effect on the number of customers contacting the company by not joining a price comparison site

This scenario will provide the company important information about potential customers and more vitally, if they will still generate the same number of customers. If company x notice a steep decline in customers, then this could mean that the company may not be as profitable as before, thus potential job losses. If the company discovers no effects, then they could proceed as before, without considering extra costs to the aggregators.

5.2.5 Customer relationship

Car insurance companies tend to give six weeks' notice of the customer's renewal price before the renewal date. Potentially, this could lead to customers searching for cheaper car insurance quotes, for a cheaper price. For a company that concentrates on new customer sales, this may have an impact on customer retention. Joining a price comparison site may introduce new customers to the company, but it may have a negative effect on loyalty (Laffey, 2009). Price comparison sites have the potential for car insurance companies to lose their loyal customers due to the ease of comparing their car insurance renewal quotes. This does not mean that the car insurance company should not try to maintain their most profitable customers. Reichheld (1996, p. 57) notes that the best way to locate such customers is to answer three questions:

1. Which of your customers are the most profitable and loyal?
2. Which customers place the greatest value on what you offer?
3. Which of your customers are worth more to you than to your competitors?

Knowing whether a profitable customer is going to defect or not can cause some concern to car insurance companies, as they will want to adjust their pricing to keep the most profitable customers. Knowing what makes a customer defect 'can point to common traits among customers who stay longer' (Reichheld and Sasser, 1990, p.109). The more the company knows about their customers, the easier it will be to try to stop them from defecting. A customer who bought their insurance directly from the insurance company, for example may be easier to retain, as they could have been influenced by the insurance brand, rather than the aggregator brand. Although to fully understand all of the different mechanisms that may make a customer leave would require big data analysis and qualitative research, within car insurance the most common reason tends to be increase in premiums (Cohen, 2012). Only when this factor has been attributed, other patterns may occur for richer customer insight.

Knowing your customers' wants and fulfilling their needs can enhance a company's profits. Satisfying a customer's needs has a positive effect on retention (Lam *et al.*, 2004). It has been shown that in many different industries a reduction in people defecting by 5 % has generated extra profits for many different companies e.g. '85% more profits in one bank's system, 50% more in an insurance brokerage and 30% more in an auto-service chain' (Reichheld and Sasser, 1990, p.107). This demonstrates that customer loyalty programs work across many different industries and that customers do respond to them positively.

Retaining profitable customers is just as important in the insurance industry as in other industries. When the insurance company embarks on a loyalty programme it has to gives the customer a sense of belonging and that they can see the relevance of the offerings to the product (Brophy, 2013). Insurance companies will need to consider the processes and their channel functions to enhance the customers' convenience and to satisfy their needs.

Customer relationship building is important, so we would need to consider the impact of customer retention, which leads us to the next scenario, again with the previous scenarios recalled for easy follow through:

Scenario 1: - the effect on marketing if a company does not join a price comparison site
Scenario 2: - the effect on the number of customers contacting the company by not joining a price comparison site
Scenario 3: - the effect of customer retention if a company does not join a price comparison site

As mentioned in chapter two, business growth can occur in two ways for company x, new acquisitions and customer retention. Scenario two focuses on customer acquisition and scenario three considers customer retention. If customers are renewing at a much smaller rate due to aggregators, then this will cause the company to consider its core strategy. Without customers, new and returning, the company should expect a decrease in profits, which would create anxiety with their shareholders.

5.2.6 Financial costs

When considering implementing a price comparison site in the business, costs will need to be considered. Aggregators will affect many areas of the insurance departments, e.g. sales, marketing and IT. The IT systems may need to be updated so that they can work with the aggregator sites and the protection of customer sensitive data. The changes within the IT department may include new staff, software and hardware so that the project can take place. To join a price comparison site the IT infrastructure needs to be informed and updated. IT infrastructure can 'account for more than 58 percent of the total IT budget of large firms' (Broadbent and Weill, 1997, p.77). This demonstrates that joining a price comparison site will affect the budgetary requirements for the company.

Not only are the costs relevant to the company, but also to the customer. When a car insurance company joins a price comparison site, the insurance company has to be aware of the switching costs, the time and effort required by the customer when changing providers (Porter, 2004). As a customer does not need to download any new software or invest in new hardware, the switching cost of using a price comparison is almost zero (Chircu and Kauffman, 2000). Aggregators tend to benefit the customer when considering switching cost strategy more than the car insurance company, but the company may reach customers that would not have gone to them initially.

Whether or not a car insurance company joins a price comparison site, the insurance company's profitability needs to be considered. The two essential factors that establish profitability are:

- Industry structure, which determines the profitability of the average competitor; and
- Sustainable competitive advantage; which allows a company to outperform the average competitor

(Porter, 2001, pp.68)

Competitors keep their future strategies a closely guarded secret, to make sure that their competitors do not beat them to it and become potential market leaders. It has been shown that established companies are more inclined to embrace a wait-and-see strategy to electronic mediation opportunities (Granados *et al*. 2013). This risk-averse approach may mean that established companies might lose out on early potential gains at a cost to their customer base. Another reason why an insurance company may not join a price comparison site is so that the insurance company does not become too reliant on them (Chung *et al*., 2012).

The price of the insurance affects the profitability of the company. The arrival of the internet and customers purchasing from the websites did not affect the premium too much as the customer needed to contact the different companies individually. The internet benefited the insurance companies more than the customer as it meant employing less staff to answer numerous phone calls for a car insurance quote. What has been shown though is that when a company is on a price comparison site, this does have a reducing effect on their premiums (Brown and Goolsbee, 2002; Baye *et al*., 2004). With the search costs lower, car insurance companies need to be aware of their competitors' quotes, and would need to at least match them, thus reducing the premiums. Lower premiums may not mean less profit though, as the company could gather more customers.

Aggregators give car insurance companies a wider basket of potential customers, but they must not neglect their core principles. Even though aggregators may give some companies some competitive advantage, total reliance on IT will not sustain this advantage (Mata *et al*. 1995). Customers may go to companies that are not on price comparison sites, if their original insurance company only concentrated on the IT side instead of the human and relationship side.

This section considers the financial costs of joining a price comparison site, which prompts us to consider scenario 4 highlighted below with previous scenario shown for easy follow-through.

Scenario 1: - the effect on marketing if a company does not join a price comparison site
Scenario 2: - the effect on the number of customers contacting the company by not joining a price comparison site
Scenario 3: - the effect of customer retention if a company does not join a price comparison site
Scenario 4: - Is it worth investing in extra resource and expenditure to enable aggregators?

As with all projects major projects in companies, they have to be cost effective. Such recent IT disaster projects where costs have spiraled and were no longer cost-effective have been the BBC digital media project that cost £100 million (Ghosh, 2013) and the £10bn abandoned NHS IT project (Syal, 2013). These two examples demonstrate the need to fully understand the complete project with costs, so that the company can still function, even if they do enable aggregators into their distribution channels.

5.3 Theoretical background on regression models

This section explores the methods taken to model the effects of price comparison sites within the UK car insurance environment, regarding sales, retention, marketing and return of investment (ROI).

Linear regression

To test the impact aggregators have had on total sales a linear regression model was constructed. A simple linear regression model (equation 5.1) is used to test if there is a relationship between a dependent variable (y) and an exploratory variable (x).

$$Y = \alpha + \beta x + \varepsilon \tag{5.1}$$

In this scenario Y is calculated by β multiplied by x, plus a constant (sometimes referred to as the intercept) α. This relationship is not always exact, so an error term needs to be introduced into the equation ε. The error term ε tends to have a zero mean and a constant standard deviation σ. The variance σ^2 (equation 5.2) is calculated by squaring the standard deviation, which calculates how far the observations are from the mean. Equation 5.1 presents provides a trend line between the dates and the total sales.

$$\sigma^2 = \frac{1}{n-1}\sum_{t-1}^{n}(y_t - \bar{y})^2 \tag{5.2}$$

where $\bar{y}$ is the mean and n is the number of observations

Time Series

The linear regression model provides a straight-line curve which does not always suffice in a time series environment. Within time series modelling the variance has to be considered. If the variance is constant and independent of x, then this is known as homoskedasticity, but if the variance varies with the size of x, it is known as heteroskedasticity. The variance is important when considering time series as it can dictate which time series technique to use.

This research uses four different time series modelling techniques to provide further insight into the sales behaviour and to forecast future sales: autoregressive (AR); moving average (MA); autoregressive moving average (ARMA); and generalized autoregressive conditional heteroskedasticity (GARCH).

Autoregressive models

Autoregressive models are a very common time series technique. Autoregressive uses the variables history with its present figure to predict the dependent variable and is usually dictated AR(p). The p represents the period up to which the historical data will be used and is usually referred to as the order of the autoregressive process.

$$Y_t = \sum_{i=1}^{p} \varphi_i \, Y_{t-i} + \varepsilon_t \qquad (5.3)$$

where $\varphi_1, \ldots, \varphi_i$ are the **parameters** of the model
Due to autoregressive models using their history, autocorrelation will need to be considered. Autocorrelation refers to the correlation of a time series with its own past.

Moving average models

Another common method in time series is the moving average model, usually dictated as MA(q) (Equation 5.4). The moving average uses the mean of the previous 'q' observations to smooth out any fluctuations.

$$MA = \frac{\sum_{i=1}^{q} \theta_i}{q} \qquad (5.4)$$

where the θ_1, ..., θ_q are the parameters of the model and q represents the order of the moving average process

Autoregressive integrated moving average model

Box and Jenkins (2008) combine the AR(p) and MA(q) techniques to produce an autoregressive integrated moving average model, usually dictated as ARIMA(*p*,*d*,*q*). These models may contains a mixture of autoregressive terms *(p)*, moving average terms *(q)*, differencing terms *(d)* or use all three. Different types of ARIMA models are shown below to demonstrate their versatility.

An ARIMA(0,1,0) model is sometimes considered as a random walk model as it concerns with the differencing aspects only

$$y_{(t)} - y_{(t-1)} = \mu \quad (5.5)$$

where μ is the constant term

If a series is said to follow a random walk, the series itself is not random, but the changes from one period to the next are. This means that the past cannot be used to predict the future.

Sometimes, using differences could lead to the terms being autocorrelated, therefore autoregressive and/or moving averages are used to fix such problems. An ARIMA(1,1,0) is a differenced first-order autoregressive model, which tend to be used if the errors of the random walk model are autocorrelated.

$$y_{(t)} = \mu + y_{(t-1)} + \emptyset(y_{(t-1)} - y_{(t-2)}) \quad (5.6)$$

where μ is the constant term and θ is the autoregressive order

An ARIMA(0,1,1) uses the moving average to correct the autocorrelated errors in the random walk.

$$y_{(t)} = \mu + y_{(t-1)} - \theta e_{(t-1)} \quad (5.7)$$

where $e_{(t-1)}$ is the error at period t-1 and θ is the coefficient of the lagged forecast error.

For a ‘mixed model’ that has all terms ARIMA (1,1,1) we have

$$y_{(t)} = \mu + \emptyset(y_{(t-1)} - y_{(t-2)}) + y_{(t-1)} - \theta e_{(t-1)} \quad (5.8)$$

Normally, an unmixed model is used because including both terms could lead to over fitting, thus making the model unsuitable for use.

Generalized autoregressive conditional heteroskedasticity model

When using ARIMA models, researchers assume constant volatility (homoscedasticity). For time series models that have a non-constant volatility (heteroscedasticity), Engle (1982) derived the Autoregressive Conditional Heteroskedasticity (ARCH) model. This model was then developed further independently by Bollersev (1986) and Taylor (1986) to produce the generalized autoregressive conditional heteroskedasticity, usually denoted as GARCH(p,q). Volatility loosely refers to the variance from the sample observations (Equation 5.2).

The GARCH *(p,q)* can be expressed as follows

$$y_t = x_t + \varepsilon_t \tag{5.9}$$

$$\varepsilon_t = \sqrt{h_t}\,.z_t \tag{5.10}$$

where z_t = is a sequence of independent, identically distributed random variables with zero mean and h_t (the variance) is expressed as

$$h_t = \omega + \beta_1 h_t \varepsilon_t^2 + \beta_2 h_t \tag{5.11}$$

The GARCH *(p,q)* uses the conditional variance and is also a linear function of its own lags. ‘The most widely used GARCH specification asserts that the best predictor of the variance in the next period is a weighted average of the long-run average variance, the variance predicted for this period, and the new information in this period that is captured by the most recent squared residual’ (Engle, 2001, p. 160). GARCH (*p,q*) models have been used for financial forecasting in the past, (Datta *et al*., 2007) so to develop one for the insurance industry would deem appropriate.

The GARCH(1,1) model is most commonly used, as it uses a normal distribution and denotes the fact that its volatility component incorporates 1 return variance term (or autoregressive lag) with 1 volatility term (or ARCH term) reads as follows:

$$h_{t+1} = \omega + \beta_1 h_t \varepsilon_t^2 + \beta_2 h_t \tag{5.12}$$

5.4 Measurement and data

5.4.1 The data

To test the different scenarios, an established UK car insurance company was used. As it had recently joined a price comparison site, its effects could be measured. Two data sets provided from the car insurance company for this analysis. The first contained dataset contained relevant information aggregated at the monthly level covering the dates between 2006 and 2008 inclusively (table 5.1). The second dataset contained relevant information aggregated at the monthly level covering the dates between 2005 and 2009 inclusively (table 5.2).

The first data set contained

- Policy Inception Month
- Direct premium
- Direct channel marketing spend
- Direct channel sales
- Direct channel cancelled policies
- Direct channel retained policies
- Aggregator premium
- Aggregator channel marketing spend
- Aggregator channel sales
- Aggregator channel cancelled policies
- Direct channel retained policies

From policy inception month, original inception date can be derived. Original inception date is the date the company first started to insure the customer. People who quoted with the insurance company and then purchased are called customers. People who quoted and did not purchase are known as prospectors. To limit the impact of returning customers appearing as a prospector, original inception date is used. If a customer first had a quote in April 2007 and purchased and then proceeded to quoted again in April 2008, this would not be used as he first purchased in April 2007.

Table 5.1: Snapshot of data set 1 (excludes aggregator information)

Policy Inception month	Direct Premium	Direct Spend	Direct Sales	Direct Renewed 1 Year	Direct Cancelled I year
Jan-06	£9,302,864	£535,711	14623	8950	5673
Feb-06	£8,354,070	£597,194	14515	8743	5772
Mar-06	£9,287,437	£637,455	18181	11194	6987
Apr-06	£8,697,362	£577,327	17176	10506	6670
May-06	£8,814,488	£494,644	17485	10568	6917
Jun-06	£8,546,649	£508,889	17169	10420	6749
Jul-06	£8,171,530	£520,002	16629	9754	6875
Aug-06	£8,176,305	£482,807	16593	9827	6766
Sep-06	£8,030,591	£570,675	17407	9960	7447
Oct-06	£8,044,128	£615,200	15934	9253	6681
Nov-06	£6,697,540	£389,417	13885	8140	5745
Dec-06	£5,168,480	£243,855	10992	6376	4616
Jan-07	£8,779,279	£685,647	13753	8147	5606
Feb-07	£7,539,943	£857,664	14019	8080	5939
Mar-07	£8,600,331	£970,998	16894	9870	7024
Apr-07	£8,031,444	£519,181	15186	8484	6702
May-07	£7,332,181	£527,901	14826	8459	6367
Jun-07	£6,304,683	£416,555	12498	7130	5368
Jul-07	£6,259,833	£296,067	11960	6927	5033
Aug-07	£4,773,418	£472,752	10170	6013	4157
Sep-07	£4,106,008	£543,292	9225	5584	3641
Oct-07	£3,870,732	£317,317	7845	5378	2467

The data was created to monitor first year renewal rates. This means that policies that were created in January 2006 will have they policy renewal and policy cancelled numbers attached to that month. The second data set provided was aggregated at a monthly level detailing:

- Month of insurance quote enquiry
- Marketing spend
- Number of direct car insurance quotes
- Number of direct car insurance Sales

- Number of Aggregator car insurance quotes
- Number of Aggregator car insurance Sales

Table 5.2: Snapshot of table (excluding later months and aggregator information)

Month	Marketing Spend	Direct quotes	Direct Sales	Direct Premium
Jan-04	£1,152,850	213223	9648	£4,126,671
Feb-04	£1,156,962	218867	13688	£5,872,840
Mar-04	£1,304,980	241085	15375	£6,566,116
Apr-04	£1,064,349	214256	14838	£6,547,028
May-04	£1,254,696	219658	17065	£7,519,854
Jun-04	£1,149,435	194767	12393	£5,467,834
Jul-04	£1,084,320	208591	12900	£5,698,880
Aug-04	£1,363,891	209871	13862	£6,168,909

These date periods are relevant for the study as they cover the dates before the company joined a price comparison site, and afterwards. The data provided by the company details that until April 2007 the company only functioned as a 'click button' tool to retrieve a quote. That is, a potential customer had to: go to the aggregator; fill in their details; get a list of quotes; click on this company banner which did not have a quote next to it; go to the car insurance company directly, and then complete their details again on the car insurance web site.

Between May 2007 and August 2007 the company was in its testing stage to make sure that its systems and infrastructure could manage with this new channel. From September 2007, the company was fully incorporated with the aggregator.

The data provided were in CSV files, which were then converted into SAS data sets for data cleansing and summarised. This data was then transferred into Excel for further analysis and graphs.

5.4.2 Measurement

The data provided was rich in detail and had only required little data cleansing. The data came from the data warehouse developed by company x. The data provided had been used for many management information reporting, thus its reliability has been paramount for the company. The data provided insight into the direct effects of price comparison sites on:

- Sales
- Retention
- Marketing
- Quotes

As an indirect measure we could hypothesise and formulate the impacts on the business itself and the strategy the business had to apply to adopt the aggregator into their business.

Quotes and Sales

To assess the effect aggregators have had within the car insurance industry, graphical analysis was produced detailing the percentage split of quotes between the aggregator and non-aggregator channel (direct).

$$\text{For Aggregator} \qquad \frac{AGG(q)}{AGG(q) + DIR(q)} \tag{5.13}$$

$$\text{For Direct quotes} \qquad \frac{DIR(q)}{AGG(q) + DIR(q)} \tag{5.14}$$

Where DIR(q) = direct quotes and AGG(q)= Aggregator quotes

Secondly, a graph detailing the number of quotes that aggregator versus direct channels achieved by month is constructed. Customers who contacted an aggregator and contacted the company directly would be counted twice. The next scenario would investigate the effect of aggregators on the company's new business sales, which is the company's primary choice of metric for customer growth size, a graph was produced investigating the total number of new sales the company had experienced. The graph will detail four different plots

1) Aggregator sales
2) Direct sales
3) Total sales (aggregator + direct)
4) A trend line to investigate the overall effect on sales

The trend line was produce using a simple linear regression (Equation 5.1). If a customer acquired a quote from more than one channel, the initial contact channel was used.

Further analysis was also conducted on the total sales to see if time series modelling was suitable for forecasting future sales. Three different modelling techniques were considered:

- AR (1)
- MA (1)
- ARIMA (1,1,0)

A graph was produced to compare the different techniques. Two graphs were produced to measure the impact of the marketing on aggregator sales. The first graph contains a double axis graph with one axis measuring the number of direct and aggregator sales, and the other axis measuring the marketing spend. The graph will potentially investigate the relationship between marketing spend with aggregator and direct sales. The second graph is to calculate the efficiency of the marketing by using a return on investment (ROI) metric.

$$\text{ROI} = \frac{\text{(Total Premium – Total marketing spend)}}{\text{Total marketing spend}} \tag{5.15}$$

Advanced forecasting methods were used for further analysis for the ROI metric. ROI behaviour has a time varying volatility, so a GARCH(1,1) model, with a lagged value of was developed, that enabled the volatility to be calculated. The results of the model were plotted upon a graph, which also included the actual figures for comparison.

The final scenario to be investigated would be the effect aggregators have on retention. First, a graph was constructed with four different plots:

1) Aggregator retention rates
2) Direct retention rates
3) Trend line of the aggregator retention rate
4) Trend line of the direct retention rate

Retention rates are calculated as

$$\text{For Aggregator} \qquad \frac{AGG(r)}{AGG(r) + AGG(nr)} \qquad (5.16)$$

$$\text{For Direct quotes} \qquad \frac{DIR(r)}{DIR(r) + DIR(nr)} \qquad (5.17)$$

where: AGG(r) = Aggregator customers renewed; AGG(nr)= Aggregator customers did not renew; DIR(r) = direct customers renewed, and DIR(nr) = direct customers not renewed

To decide whether the customer belonged to an aggregator or direct channel route, the customer's initial quote channel was used. For the trend lines, a simple linear equation will be used as demonstrated with equation 5.3. The problem when considering a simple linear model is the assumption that all the error terms, when squared, is the same at any given point. When this is not the case then the model is said to suffer from heteroskedasticity. A GARCH(1,1) model with a lagged value of 1 was developed and plotted against actual data to verify its forecasting accuracy.

When analysing whether or not a customer has renewed, the customer will have to have had their policy for the full term, and it will have had to have been due for renewal. If the customer prior to their renewal date cancelled any policies, then these policies were excluded for final analysis.

5.5 Results

To investigate the different scenarios different graphs are produced highlighting key components that are affected by the aggregators:

- Quotes
- Sales
- Marketing
- Customer Retention

Figure 5.3: % quote split between direct and aggregator channel

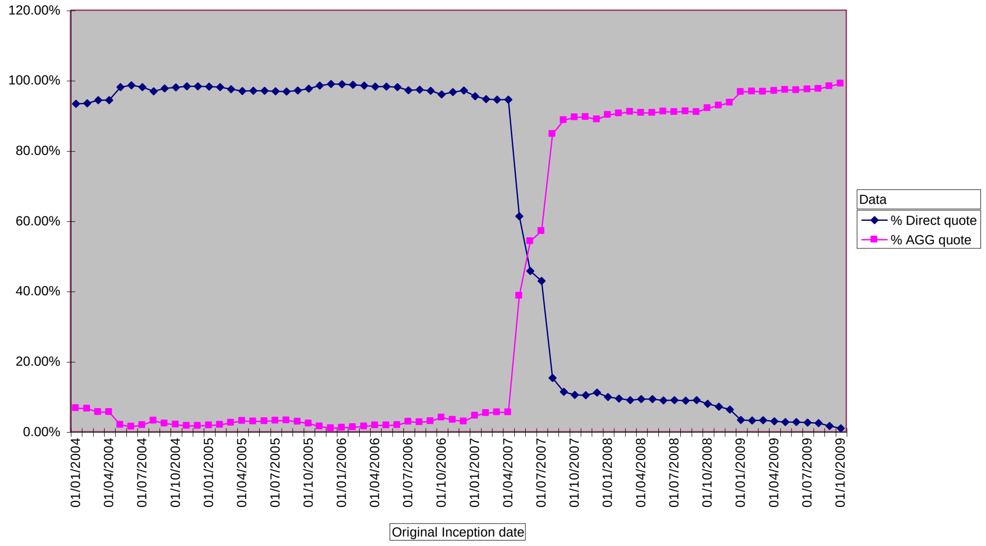
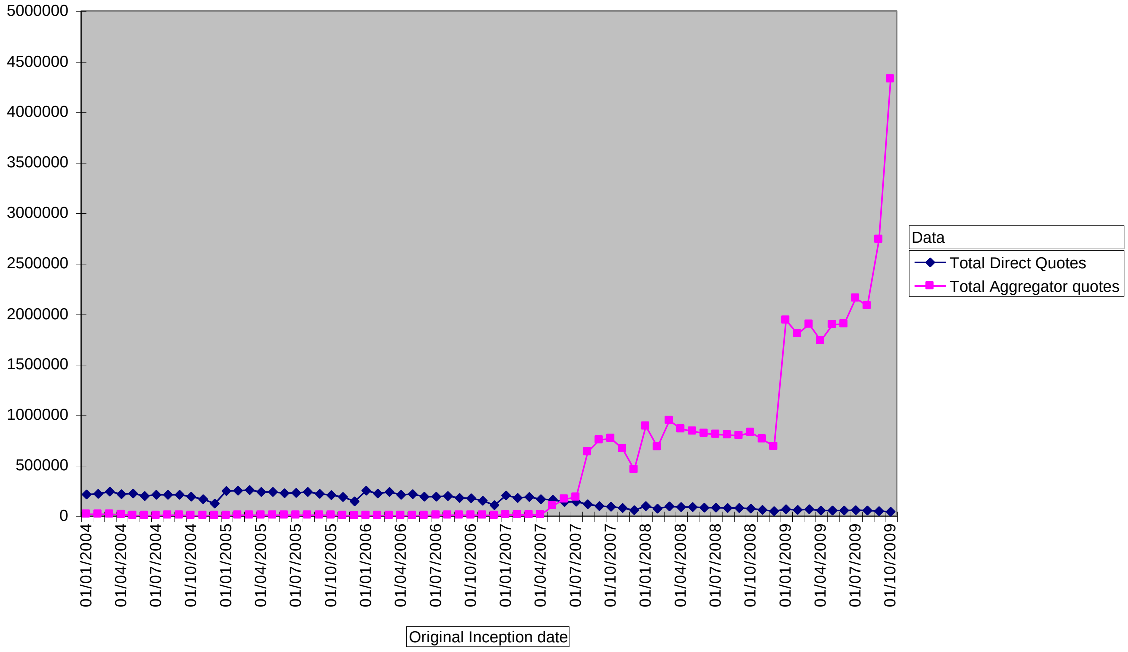


Figure 5.4: total number of quotes split by direct and aggregator channel



The results for all periods are summarized in table 5.3 below.

Table 5.3: Descriptive statistics of aggregator and direct quotes for different periods

Period / statistics	**mean**	**Std. dev.**	**max**	**min**	**skewness**	**kurtosis**
% direct quote (Jan04–Jun07)	97.05%	0.02%	99.00%	93.34%	-1.02	0.08
% agg quote (Jan04–Jun07)	2.95%	0.15%	6.66%	1.00%	1.02	0.08
Direct quote (Jan04–Jun07)	202,891	34,345	254,541	107,140	-0.86	0.68
Agg quote (Jan04–Jun07)	6,209	3552	15,271	1,452	1.19	0.88
% direct quote (Oct07–Sep09)	6.70%	0.03%	11.12%	1.65%	-0.32	-1.70
% agg quote (Oct07–Sep09)	93.30%	0.03%	98.35%	88.88%	0.32	-1.70
Direct quote (Oct07–Sep09)	68,988	15,413	97,039	46,015	0.19	-1.19
Agg quote (Oct07–Sep09)	1,238,268	644,561	2,736,920	456,796	0.78	-0.76

Figure 5.3 and table 5.3 demonstrates the proportion of quotes that arrive to the insurance company via the price comparison site and answers scenario 2, the effect on the number of customers contacting the company by not joining a price comparison site. Figure 5.3 also illustrates how rapidly a price comparison site can dominate a customer's choice of channel. Figure 5.4 demonstrates the actual figures and the extra coverage aggregators provide an insurance company. The shape of the distribution of the aggregator quotes can be considered exponential, thus demonstrating the input aggregators can have for the business. In 2009 direct channels had a total of 547,083 quotes, which is less than half the amount of quotes achieved by an aggregator in one month (January 2009 received 1,936,986). The insurance company cannot directly influence customers to go to a certain price comparison site: that will always remain within the remit of the aggregator and their marketing team/budget.

Figures 5.3 and 5.4 demonstrate that aggregators can increase the amount of potential customers gathering a quote from the company, but the actual number of sales needs to be investigated.

Sales

Figure 5.5: Total sales by month split by aggregator and direct channel

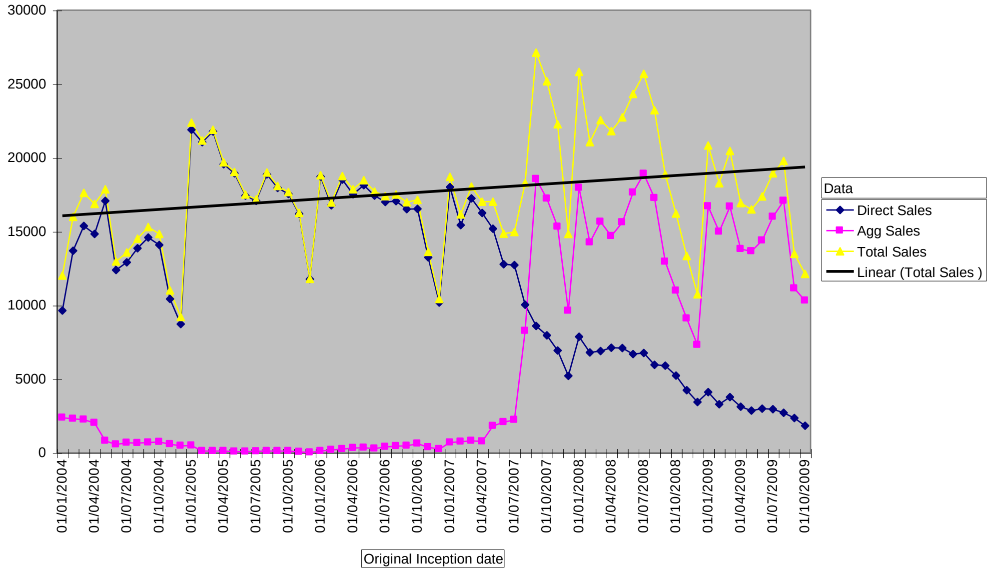


Table 5.4: Descriptive statistics of aggregator and direct sales for different periods

Period / statistics	**mean**	**Std. dev.**	**max**	**min**	**skewness**	**kurtosis**
Direct sales (Jan04–Jun07)	16,055	3,159	21,897	8,730	-0.50	0.00
Agg sales (Jan04–Jun07)	561	615	2,364	24	2.01	3.54
Total sales (Jan04–Jun07)	16,620	2,972	22,382	9,205	-0.64	0.36
Direct sales (Oct07–Sep09)	5,093	1,875	7,954	2,362	0.01	-1.62
Agg sales (Oct07–Sep09)	14,537	3,026	18,911	7,302	-0.88	0.13
Total sales (Oct07–Sep09)	19,630	4,120	25,822	10,747	-0.36	-0.55

Table 5.5: Correlation statistics between direct sales and aggregator sales Jan 04-Jun 07

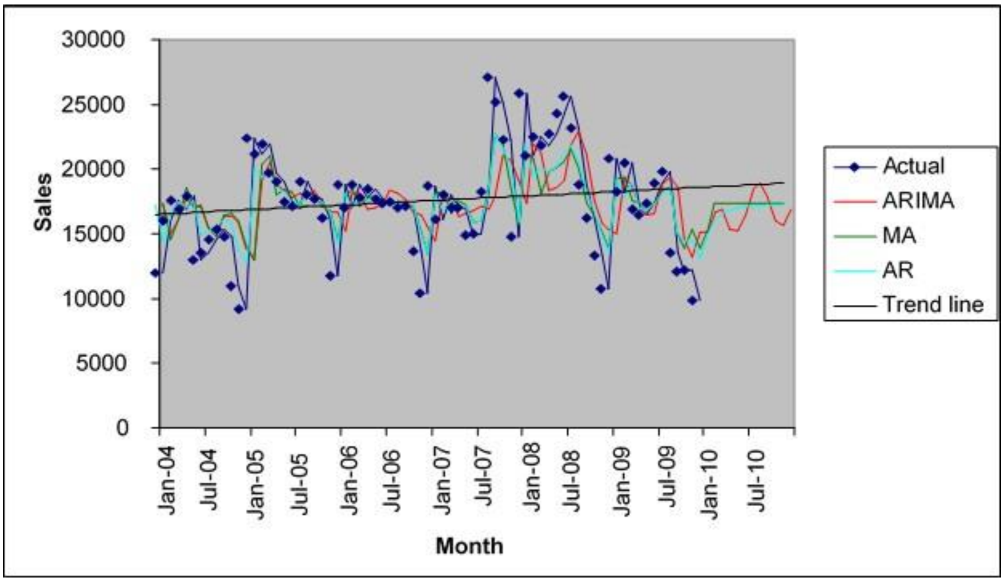
	Direct sales	***Aggregator sale***
Direct sales	*1*	*-0.392 (0.012)*
Aggregator sale	*-0.392 (0.012)*	*1*

Table 5.6: Correlation statistics between direct sales and aggregator sales Oct 07-Sep 09

	Direct sales	***Aggregator sale***
Direct sales	*1*	*0.379 (0.067)*
Aggregator sale	*0.379 (0.067)*	*1*

From Figure 5.5, it can be shown that prior to the company fully adopting the aggregator into its channels mix, the direct channel had started a downward trend. This pattern is even more extreme when the company is fully integrated with the aggregator. What is prevalent are the number of sales via the aggregator channel not only makes up for the shortfall of direct sales, but also increases the overall customers. Also from Figure 5.5 the trend line shows that the extra sales generated from the aggregators have reversed the sales decline. Finally, figures from tables 5.5 and 5.6 shows that there is no correlation between direct sales and aggregator sales. From these statistics and graphs we can partially answer scenario 4, is it worth investing in extra resource and expenditure to enable aggregators. Without aggregators, we can presume that the number of sales would continue to drop, so for maintained customer growth, aggregators could provide the answer.

Figure 5.6: Comparing different time series techniques for sales



From Figure 5.6, the plots of any of the three time series techniques could be used, but for forecasting, the ARIMA(1,1,0) plot tends to provide further insight. The plot does demonstrate that even though a trend line can provide some insight, to get a clearer view of the sales pattern, autoregressive moving average time series analysis would need to be considered for more accurate monthly forecasting.

Figure 5.7: Marketing spend and aggregator sales

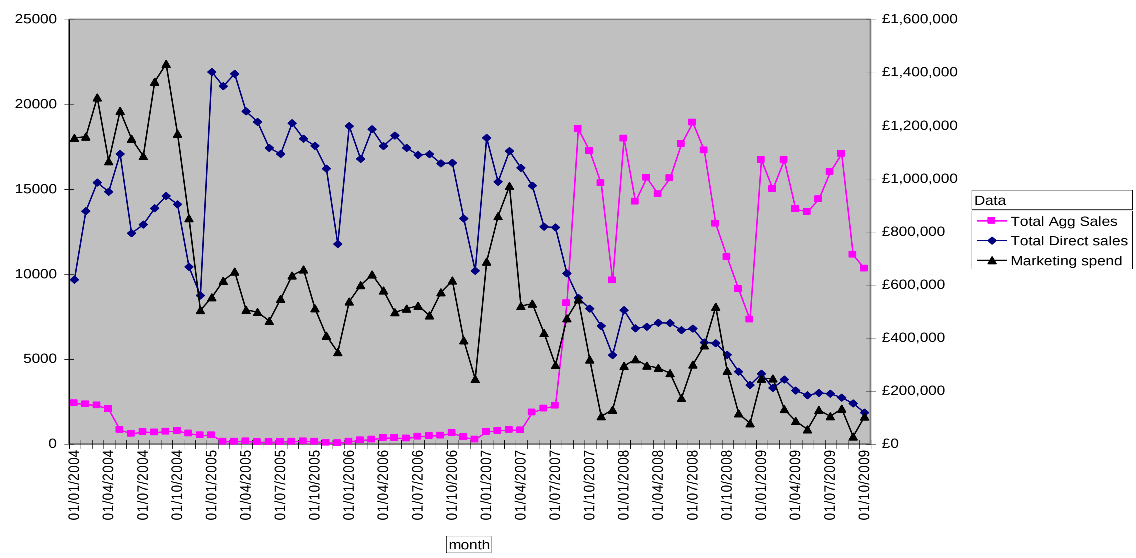


Table 5.7: Descriptive statistics of marketing spend for different periods

Period / statistics	**mean**	**Std. dev.**	**max**	**min**	**skewness**	**kurtosis**
Marketing spend (Jan04–Jun07)	£726,610	£317,032	£1,431,977	£243,855	0.83	-0.57
Marketing spend (Oct07-Sep09)	£207,085	£119,852	£515,638	£27,022	0.58	0.10

Table 5.8: Correlation statistics between direct sales, aggregator sales and marketing spend Jan 04-Jun 07

	Direct sales	***Aggregator sale***	***Marketing spend***
Direct sales	*1*	*-0.392 (0.012)*	*-0.274 (0.087)*
Aggregator sale	*-0.392 (0.012)*	*1*	*0.670 (<0.001)*
Marketing spend	*-0.274 (0.087)*	*0.670 (<0.001)*	*1*

Table 5.9: Correlation statistics between direct sales, aggregator sales and marketing spend Oct07-Sep09

	Direct sales	***Aggregator sale***	***Marketing spend***
Direct sales	*1*	*0.379 (0.067)*	*0.659 (0.005)*
Aggregator sale	*0.379 (0.067)*	*1*	*0.352 (0.091)*
Marketing spend	*0.659 (0.005)*	*0.352 (0.091)*	*1*

Table 5.10: Descriptive statistics of aggregator and direct sales for different periods

Period / statistics	**mean**	**Std. dev.**	**max**	**Min**	**skewness**	**kurtosis**
ROI direct (Jan06–Jun07)	114.0%	3.17%	120.2%	108.8%	-0.441	0.742
ROI agg (Jan06–Jun07)	111.7%	1.32%	117.7%	109.9%	1.655	2.383
ROI all (Jan06–Jun07)	114.1%	3.14%	119.8%	107.8%	-0.578	0.558
ROI direct (Oct07–Sep09)	115.2%	9.50%	147.2%	104.8%	1.943	4.611
ROI agg (Oct07–Sep09)	109.4%	0.41%	110.5%	108.7%	0.762	0.611
ROI all (Oct07–Sep09)	110.1%	1.38%	107.1%	112.9%	0.022	0.047

The company for this research uses the marketing budget to pay for their price comparison sales. Figure 5.7 demonstrates that a lower marketing spend and an introduction of aggregators may also contribute to lower direct sales. The pattern of aggregator sales follows the pattern of marketing spends, which suggests the branding effect. From tables 5.8 and 5.9, it was shown that a correlation occurs between marketing spend and aggregator sales pre integration but not post integration. Also, before aggregator integration, marketing spend on direct sales were not correlated, but they are post integration. This demonstrates that when the case company x has a link on an aggregator website, the advertising was strong enough to make customers complete the quote process again. Finally, from tables 5.8 and 5.9, marketing does influence direct sales in an aggregator environment.

Figure 5.8: Return of investment (ROI) split by direct and aggregator channel

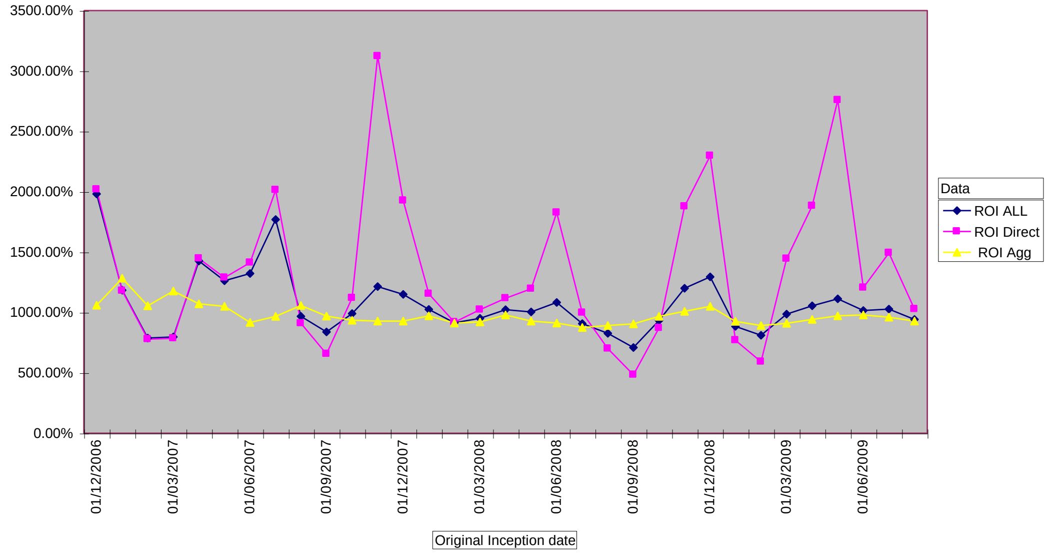
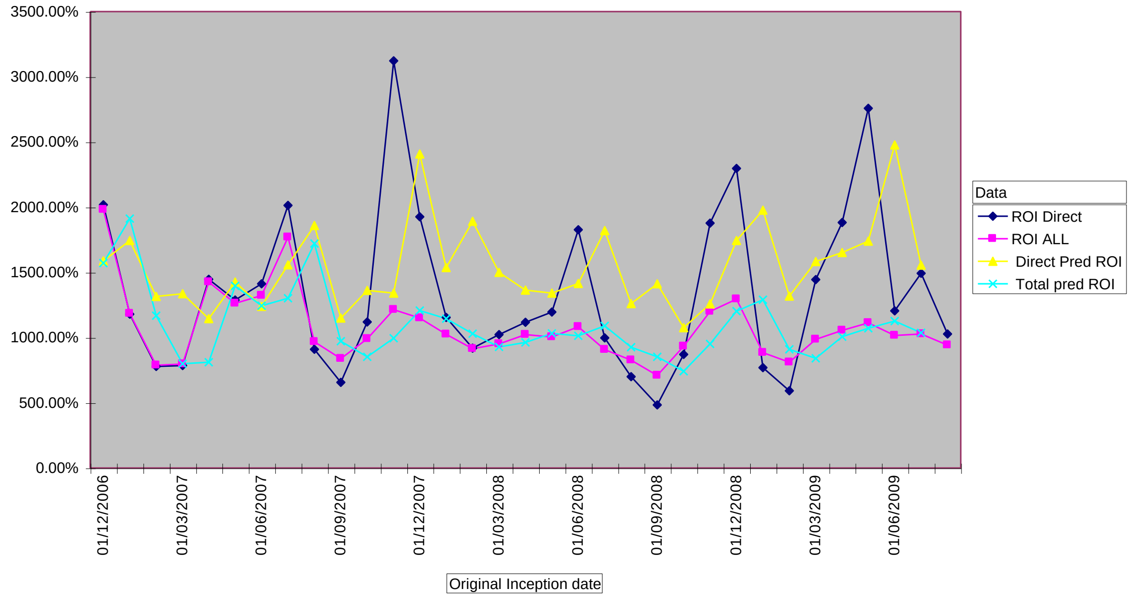


Figure 5.9: Predicted ROI rate v actual for direct and all



From Figure 5.8, as expected, the ROI ALL plot was historically influenced by 'direct', but is now mainly influenced by the aggregator due to the majority of the business coming from this channel. The ROI used is a premium-based, not a value-based metric. The direct channel does produce a better ROI when considering premiums only, but aggregators provide more sales, thus potentially facilitating company growth. Figure 5.8 also demonstrates that the ROI for aggregators is stable, with little fluctuation across the mean, thus should not be considered with a GARCH(1,1) model. From table 5.10, it can be shown that aggregators have a slight negative effect on the total ROI (114% to 110%). This is predominantly driven by the lower ROI attributed to the aggregators.

Figure 5.9 compares the GARCH(1,1) model of Direct ROI and All ROI against actual ROI. The graph demonstrates that ROI behaves in a stochastic manner and that GARCH(1,1) could be considered for forecasting ROI.

Figure 5.10: Retention rates split by direct and aggregator channel

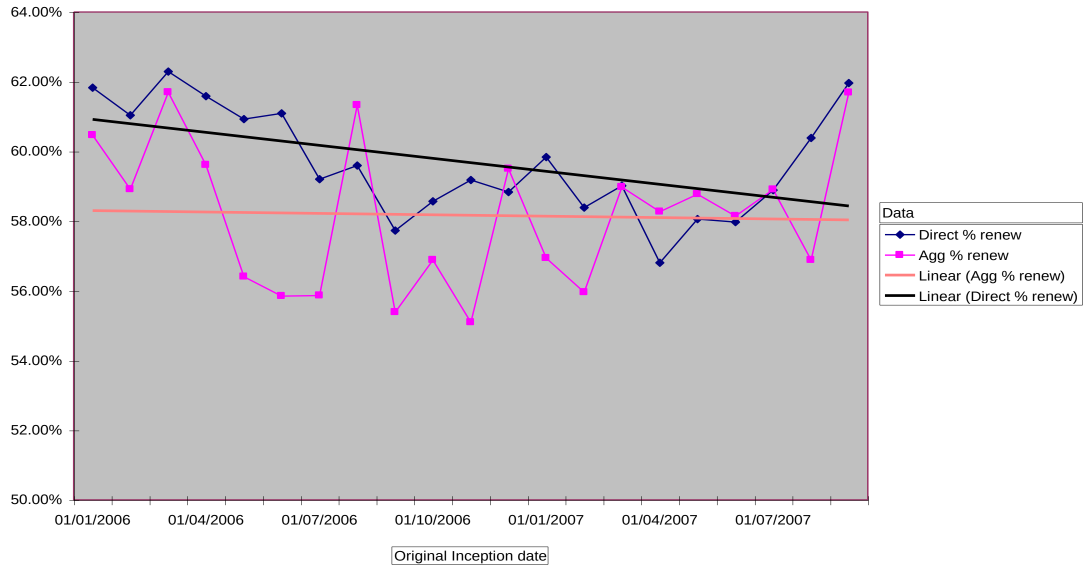


Table 5.11: Descriptive statistics of aggregator and direct renewal rates

Period / statistics	**mean**	**Std. dev.**	**max**	**min**	**skewness**	**kurtosis**
Direct renewal	60.1%	0.03%	70.2%	56.8%	-2.48	8.73
Agg renewal	58.9%	0.04%	74.2%	55.1%	2.86	10.65

Table 5.12: Correlation statistics between direct and aggregator renewal rates

	Direct renewal	***Aggregator renewal***
Direct renewal	*1*	*0.833 (<0.001)*
Aggregator renewal	*0.833 (<0.001)*	*1*

From Figure 5.10 and table 5.11, scenario 3, the effect on customer retention can be observed. The figures show that renewal rates have stayed consistent with aggregators, but have been decreasing for customers who contact the business directly. The reduction in direct channel retention rates suggests that customers may be more willing to research their renewal quote, thus leading more such customers to aggregators. From table 5.12, it shows that the renewal rates between direct channels and aggregators channels are highly correlated. This gives rise to suggest that renewal rates may be driven by additional factors and is not purely channel specific, which is outside the scope of this research.

Figure 5.11: Predictive and actual retention rates

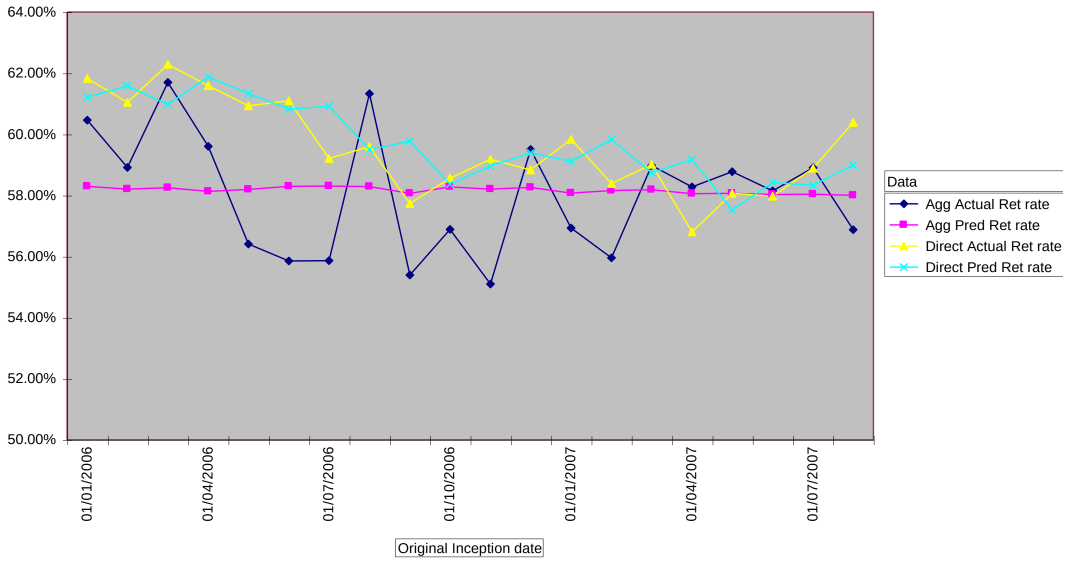


Figure 5.11 demonstrates that direct renewal rates behave in a GARCH(1,1) manner, unlike aggregators. From Figures 5.10 and 5.11, aggregator renewal rates may produce the same error around the mean, thus making it difficult to impose a GARCH(1,1) model to predict it. From viewing the figure 5.11, it can be observed that the aggregator renewal rate seems erratic, with its peaks and troughs, compared to the actual predicted rate. As aggregators are relatively new there is limited information available, so it would be difficult to know if this will be the aggregator renewal rate long term behaviour or not.

5.6 Summary and conclusion

5.6.1 Summary

Figures 5.3 to 5.11 and tables 5.4 to 5.12 provide insight into a company's behaviour in an aggregator environment. Considering the four different scenarios of a company not joining a price comparison site, the following table was constructed

Table 5.13: Results of scenarios

Scenario	**Description**	**Results**
1	Effect on Marketing	A Good Return on Investment
2	Customer Contacting	Reduction in new business
3	Customer Retention	Reduction in customer Retention
4	Extra resource/expense	Increased sales will make this feasible

From the graphs, it can be shown that time series modelling can be used in a number of scenarios within the car insurance industry. Although sales can use simple time series techniques, for ROI and retention rates, more advanced modelling techniques need to be considered.

5.6.2 Discussion

The findings provide insights into the UK car insurance market within an aggregator environment, by producing a new business model and investigating alternative scenarios.

For scenario 1, Figures 5.8 and 5.9 with table 5.12 have shown we would expect marketing to produce a stable return of investment if the company did not join an aggregator site as time progressed. The number of new business sales had been declining prior to the company joining the aggregator, as more potential customers are being driven to price comparison sites. If the company chose to reduce its marketing to try and improve its ROI, this may lead to more potential customers not contacting them directly, thus being counter-intuitive. Also, the graphs had shown that marketing may influence customers' choice on price comparison sites, but table 5.9 showed this effect to be weak with a correlation statistic of 0.352. The results also show that aggregators are taking customers away from contacting the company directly. The results correspond with the findings of Coelho and Easingwood (2003, p.32) in that price comparison sites 'generate a source of variable costs and provide a much faster and wider market coverage'.

From Figure 5.7, it is also noted that the direct ROI is greater than the aggregator ROI, which could be caused by price transparency. The reduction in ROI could be due to the way that aggregators display all the available prices, which may narrow price discrepancy, which follows the findings of Stigler (1961).

For scenario 2, we could see from Figures 5.3 and 5.4 and table 5.4 that if a company decides not to join a price comparison site, then we would expect fewer people to contact the company. Prior to joining an aggregator, the number of customers contacting the company for a quote had been decreasing, even though the marketing budget had stayed quite consistent. By employing a multi-channel strategy, the results generated more customers, supporting Neslin *et al*. (2006) and Blattberg *et al*. (2008). The extra customers may have been gained from lesser known brands (Leuthesser *et al*., 1995), but this would need further research.

As aggregators tend to have a bigger marketing budget than car insurance companies, the channel choice of the customers tends to be the aggregator site, which is consistent with Ansari *et al* (2008).

Scenario 3 considers retention (Figure 5.10 and table 5.11), price comparison sites have a lower rate of renewal then direct channels, but the trend of the direct channels was decreasing. When switching costs are set higher it is easier to retain customers (Gronhaug and Gilly, 1991), but with reduced switching costs, the company should expect a lower retention rate. The lower retention rate with aggregators could be due to the fact that they are a web based tool (Ansari *et al*., 2008).

Another issue that would need to be considered with the renewal rates is the relationship building process with the customer. If the first contact with the insurance company is via an aggregator and the person then purchases via the aggregator, then this limits the possibilities of building that initial relationship with the customer, which is consistent with Coulter and Coulter (2002)'s research.

Finally, scenario 4, as all the extra quotes are completed online the infrastructure costs need to be considered in order to ensure all customers receive a quote. Figure 5.3 and table 5.3 demonstrates that aggregators do take customers away from contacting the company directly, which corresponds with the research conducted by Granados *et al*. (2008). The number of quotes the company achieves does grow significantly, so the IT systems will also need to be updated, even if only to ensure that the quote reference ID's do not exceed the current company limit of unique references. This highlights one of the issues of why infrastructure is so important and can use 60% of the budget requirements (Broadbent and Weill, 1997).

For insurance, companies that notice that their sales and retention are decreasing, price comparison sites could potentially provide a solution, but they come at a cost. As shown from the business model and analysis, for a company to encompass aggregators not only can they not influence the type of customer who goes there, they may also notice a decrease in customers contacting them directly. This scenario may lead to the price comparison site gaining more influence on the insurance company. What a company can gain from joining an aggregator is a wider exposure to the customers they usually do not reach but want, with a possibility of an increase in sales. Price comparison sites could also reduce the need for extra staff to cover potential bigger sales volumes, as all the processes can be carried out online.

The findings show that GARCH(1,1) is indeed a suitable specification for describing return and risk behaviour of the insurance companies.

5.6.2 Conclusions

Aggregators impact insurance companies on all the different mechanisms of the business model. If an insurance company does not join a price comparison site, then it can expect its sales and size to reduce. Joining a price comparison site requires a change in the business model to encompass the extra complexities of a new channel; IT infrastructure to deal with the numerous additional quotes and provide accurate information to the price comparison site; staff training, and a change in marketing.

Joining a price comparison site may not provide a 'quick fix' for a company that is noticing its sales volumes dropping, but could provide further insight into their current strategies. The different scenarios contribute to the development of the IBRM as shown in figure 5.12.

Figure 5.12: Developing the Integrated Business Reintermediation Model (part 2)

- Marketing does influence direct sales in an aggregator environment
- Aggregators will reduce the declining customer figures

- Increased sales will make this feasible
- Infrastructure will need improving to handle increased quote volumes

Value Proposition Product/Price Marketing the product/brand Distribution channels	**Value Relationship** Value chain Relationship building with partners Infrastructure
Customer Relationship Locating profitable customers Creating relationship with customers Satisfying customer needs	**Financial Costs** Costs of the other segments Profitability Competitor cost strategy

- aggregators are taking customers away from contacting the company directly
- price comparison sites customers have a lower rate of renewal then direct channels

- ROI is positive after integrating with aggregators
- Aggregators introduce new customers to the brand

5.6.3 Further work

This chapter provides a foundation for further research into the field of price comparison sites and the car insurance industry. It could be extended to review a longer time period of retention rate to see if the retention for direct channels keeps on decreasing and to investigate practices to increase customer retention. Research into the purchasing habits during the car renewal period, for example whether a customer uses aggregators, would also provide some interesting findings.

The results also provide insight into customer behaviour and renewal rates, return of investment and sales, using time series and stochastic models. This area could be extended to model the percentage quote split between direct and Aggregator Channel. This pattern sees a dramatic shift between both channels which would provide some further insights into implementing a price comparison site into the acquisition channel mix.

Although the ROI provides an initial insight into the pricing structures of the company, if claims data was available then value analysis could also have been incorporated, as even though price comparison site have a lower ROI, they may produce a higher customer value. In this research, the company did see a rise in customers when they joined a price comparison site, but without knowing at what cost, e.g. lower retention rates, new drivers or poaching customers from competitors.

Finally, all of this research is based on a company joining an aggregator, so research into a company that does not join an aggregator would provide a balance. In this vein, comparison of the prospects of the case company before and after joining and aggregator provides vital clues to the effects of not joining.

Chapter 6: Customer segmentation in an aggregator environment

6.1 Introduction

Customer relationship management (CRM) has the power to generate extra value for the company as well as its customers. The UK car insurance environment has gone through some turbulence with the introduction of price comparison sites (aggregators). This chapter uses this new environment as the premise of this research to promote a powerful segmentation technique based on predicted customer value, actual customer value and predicted renewal rate. The predicted value model used a different modelling and data mining techniques and as well as Winsorization on the data source. Winsorization is a robust regression technique designed to control the influence of outliers by setting extreme values a specified percentile. For all the different data mining and modelling techniques, quantile regression produced the most accurate model, which, coincidently, is the only modelling technique in this research that is unaffected by Winsorization. These results will have important strategic repercussions for the UK car insurance industry, by way of focussing their customer retention plans on profitable customers.

The arrival of price comparison sites (aggregators) has changed the shopping habits of people purchasing car insurance. Price comparison sites have made it easier for customers to compare numerous car insurance companies by using one internet site, instead of approaching each company individually. Aggregators make their money every time someone purchases their car insurance via their website, so they would want customers to leave their current car insurers, at a cost to the car insurance company. It is generally known that keeping new customers is more cost effective than acquiring new customers. This has led to the creation of customer relationship management (CRM), where targeting the most profitable customer is commonplace.

The literature for CRM within the insurance industry tends to include customer lifetime value (CLV) for its customer segmentation. Previous research has used potential value as a means to target customers who could generate the most profit for the insurance company (Verhoef and Donkers, 2001; Ryals and Knox, 2005; Guillén *et al.*, 2011). Data mining tools have been used within the car insurance industry with Smith *et al*. (2000) using them for retention and to predict the likelihood of a customer making a claim. The most common absentee from this research is the UK car insurance.

Customer campaign modelling is used for customer relationship management (CRM) to help reduce churn and increase their profit from the customer. One of the most common customer segmentation techniques involves customer lifetime value (CLV). Customer lifetime value is important to companies as if it is used correctly, it can lead to more intelligent campaigns. For customers with a large CLV, it may be preferable to use a more valuable promotion to keep their business. Insurance CLV needs to be considered differently from catalogue/product CLV, as with insurance there is potentially a risk of the customer costing the company a great amount of money through claiming.

This study considers customer segmentation techniques which go beyond the use of statistical models and data mining, but also consider the outliers that appear on the dataset. The models developed are more relevant to the prediction of the subsequent period customer behaviour, rather than to the predicted CLV, so CLV is deduced from the customers predicted performance.

As mentioned above, the rationale for this study is that there is a dearth of such quantitative modelling in the car insurance industry particularly targeted on understanding, the effects of aggregators on customer retention. To the researcher's knowledge, this is the first study to explore these effects in the UK car insurance industry

The remainder of this study is organised as follows. Section 6.2 reviews the general literature on CRM. Section 6.3 discusses the theoretical background to the statistical models and decision trees used in implementing the methodology for this study. Section 4 describes the data used in this study. Section 6.5 presents the empirical analysis and the interpretation of the modelling results with implications for marketing action. Section 6.6 summarises the main results and concludes the study.

6.2 Review

From chapters 4 and 5 customer retention has been measured and investigated. This chapter expands the previous work into a framework that can be implemented within the car insurance industry. Hence, this review will delve into the mechanics of retaining the most profitable customers.

6.2.3 CRM

Companies that concentrate on building their company size need to think beyond just new acquisitions. It has been shown that 'customer defection has a surprisingly powerful impact on the bottom line' (Reichheld and Sasser, 1990 pp.105), as the cost of acquiring a new customer can be as much as five times the cost of retaining an existing one (Pfeifer 2005). Ignoring customer retention could lead to an increased cost in building the company's customer base by just focusing on acquisition only. To maintain customers through retention, customer relationship management (CRM) strategies are applied. Customer Relationship Management (CRM) can be defined as:

> *Business strategy and mode of operation deployed to maintain and develop relationships with profitable customers, and manage the cost of doing business with less profitable customers.*
>
> Stone and Foss, 2002, p.14

Insurance companies, like any other industry, can grow their customer base by taking customers from their competitors, so retaining their customers is important. Customer retention is therefore important, and has become a major factor as 'the competitive nature of the insurance industry continues to evolve … and the importance of relationship marketing practices and customer retention continues to grow' (Taylor, 2001, pp.32). Increasing customer loyalty practices should be as standard in all industries along with a clear indication of which customers it would be profitable to retain.

When a company decides to embark on a customer retention strategy, they will need to consider who to maintain, as not all customers are equal (Peppers and Rogers, 1998). Certain customers will generate more money for a company than others and within insurance this is no exception. Where the insurance industry differs from other industries is that its losses could go beyond just marketing costs. Kim *et al* (2006) note the costs of customers not fulfilling their payments on a monthly basis, but ultimately it is claims costs that demonstrate whether or not a customer is truly profitable. ‘The core value proposition of the insurance industry is risk control and risk financing. At the same time these are the two core competencies that are of utmost value to an insurance company’ (Muller & Zimmerman, 2003, pp.3). The risk aspect within insurance is balanced by the quotes given by the company. Insurance needs to consider any potential claims made against the company which in some, although rare, cases may exceed £1,000,000.

The advent of aggregators has meant that car insurance companies within the UK have had to adapt their customer retention programs. With price-to-price comparison sites (aggregators), customers who used the internet to compare prices incurred low search costs (Verhoef and Donkers, 2005). The traditional search process still involved the customers going to numerous different web sites to input their data, which was possibly preferable to phoning the insurance company, but aggregators have reduced this search cost even further. A customer can easily compare their insurance quote by visiting just one web site, instead of contacting the different insurance companies directly. This could lead to customers leaving, if they do not feel valued and think they can get a better deal elsewhere. The difficulties of potential huge claim costs along with car insurance quotes being compared so easily has made CRM practices more relevant, especially in terms of which customers it is in the company’s best interest to keep.

In insurance the customer with the lowest premium will not initially generate the biggest profit. The reason to improve retention is that the ‘profit earned from each individual customer grows as the customer stays with the company’ (Reicheld, 2001, p.37). The profit generated by the retention of low premium customers needs to be balanced against the risk of someone making a claim - this is usually carried out through the medium of a pricing model. This means that if the company does not want to keep a particular type of customer and wants to inflate their renewal price, this can cause problems with marketing. A clear line of communication must be established between the pricing and marketing departments, so that marketing stops targeting customers that the pricing department thinks are bad risks.

A customer’s reaction to marketing or the channel they use to contact a company can affect their behaviour, for example acquisition channels affect retention and customer value, but ’much of the heterogeneity across the various channels disappear[s] after the first year’ (Verhoef and Donkers, 2005, p.40). The more price-sensitive customer may leave after the first year of purchasing car insurance, so in this instance it may seem better practice for the insurance company to wait until the first year renewal cycle has completed before embarking on a CRM strategy. Knowing how the customer contacted the company may be of benefit to the company- for example a customer who contacted the company via the web could arguably be more inclined to respond favourably to an email promotion.

6.2.2 Recency frequency monetary (RFM) methodology

Companies, regardless of their industry, soon realise that not all customers are equal and that certain customers may need to be treated differently. Applying a customer segmentation strategy enables companies to group customers based on certain attributes in which they may be similar. Analysing the 'different customer segments allows marketing spend to be proportioned to deliver maximum return on investment' (Gee *et al*., 2008, p. 370). For a company to understand their customers, they have to know their data and be sure that this data is accurate. Only by knowing the data can the company produce appropriate and logical customer segmentation.

Recency, frequency and monetary (RFM) methodology is a very simple and powerful segmentation technique. Recency details how long ago the customer had bought from the company; frequency, how many times they purchase from the company in a given time; and monetary represents how much money they have spent with the company. Using RFM for marketing is not new and it has been used for over 30 years (Sohrabi and Khanlari, 2007). RFM models are used to represent the behaviour attributes associated with the customer (Chan, 2005). This implies that the more the company knows about the customer, the more insightful the RFM segmentation.

Within the insurance industry, especially within the UK car insurance industry, RFM may not seem the appropriate segmentation approach as insurance is purchased on annual basis, so frequency will be consistent amongst the majority of the customer. Yet, RFM has been shown to be useful in the insurance industry 'to help identify valuable customers and develop effective marketing strategy' (Wei *et al*., 2010, p.4205). When considering the different offers a company may use in targeting its customers, it is worth considering that the simpler the segmentation the easier it would be to implement.

RFM is a simple segmentation technique which can nevertheless be modified for different purposes. Donkers *et al*. (2007, pp.173) derived the following RFM variables: 'purchase recency dummy and a cancellation recency dummy; ownership dummies for each service; customer profitability; and relationship age.' The adaptability of the RFM segmentation, with its ability to provide more insight into a customer's relationship with the company, can also lead to a more complex segmentation technique. For an insurance company that only provides car insurance, the RFM segmentation would need adapting.

Recency could dictate how long someone has been with the company, especially for first year customers. A special welcome email should be used, with more introductory emails during the first year. After the first year, the Recency 'R' should no longer be considered and frequency should be applied, or the number of times the customer has purchased their insurance with the company. Monetary would need to consider using actual value, predicted value or both.

6.2.4 Customer lifetime value

Customer lifetime value (CLV) is a powerful segmentation technique for companies to focus on which customers it would be more beneficial to retain. Focussing on customers who may produce the greatest value for the company and not contacting everyone is a view contradictory to Reicheld (1996) (Ryala and Knox, 2005). If an insurance company decides to focus on high premium customers, this could be counterintuitive as these customers tend to be high-risk and could generate a loss to the company due to their claims costs, therefore a long-term strategy should be implemented.

CLV can be defined as the 'present value of the future cash flows attributed to the customer relationship' (Pfeifer *et al*., 2005, pp.10). This variable is therefore determined for insurance if we know the probability of a customer leaving the company (Guelman *et al*., 2012) and therefore utilising the company's historical data could prove beneficial when calculating CLV. As mentioned previously, customer value within insurance needs to consider claims cost, as some customers can have a serious negative effect on the business and should be considered for CLV (Ryals and Knox, 2005; Hawkes, 2000; Guillén *et al*., 2011).

Different statistical techniques have been used to predict customer value. The main goal for these statistical models is to find 'who are profitable customers?' (Wu *et al*., 2009, p.4). Statistical regression analysis based on the mean (Verhoef and Donkers, 2001), regression based on the median (Benoit and Poel, 2009) and data mining tools (Wang *et al*., 2005; Shen and Chuang, 2009) have been used to predict CLV, but research on comparing these different techniques is still new.

For companies not familiar with CLV, a cost ratio technique (cost of claim by total premium) could be applied (Smith *et al*., 2000). The main implication, Smith *et al*. (2000) noted, is the willingness of the company to use data mining tools. For the company this would provide another useful tool in their armoury, but without knowledge and an acceptance for data mining, they may well turn away from the idea.

6.2.5 CRM measurement

As with all projects, the result of the CRM needs to be measured, since without appropriate measurement, it would be difficult to judge the benefits of CRM. The company would need to look into their future applications, instead of their infrastructure, to monitor their results (Goodhue *et al*., 2002). If the company was to focus solely on the infrastructure, the benefits of the CRM may be overlooked when considering customer retention and recommended new acquisitions from satisfied customers informing their friends.

A company cannot judge their own CRM success by comparing against different industries. Within insurance, car insurance cannot even be compared against other types of insurance, as when compared against other insurance types the retention rates tend to be more negative (Verhoef and Donkers, 2005). This scenario enables the company to produce a less complicated analysis as they should just need to concentrate on their own customers, instead of other industry standards.

When a CRM strategy has been implemented, the company should not expect an unrealistic turnaround of retention. No company can satisfy the needs of all customers and 'customers who are worth more to a competitor will eventually defect' (Reichheld, 1996, p. 6). If the company sets an unrealistic retention rate, this could lead to the CRM project being classified a failure. Time scales must also be considered, especially if statistical models have been used to develop CLV. The company would need to ask themselves 'does customer lifetime value… turn out to be as predicted?' (Rust *et al*., 2004, p.84). This demonstrates that a long-term strategy needs to be considered when considering CLV to ensure that the models perform efficiently.

As well as customer performance, company performance needs to be considered. Although company performance may not be so easily measurable, Doyle (1989, pp.78) notes it is 'not what the producer puts in, but what the customer gets out'. If the company does not perform adequately, then they should not expect customer value to increase. This performance is expected to be enhanced by effective customer marketing and service strategies which insights from statistical analyses in this chapter support.

6.3 Theoretical background on regression models and decision trees

The section describes the different modelling and data mining tools used for customer value and retention. Different methods were used so that a comparison could be conducted to discover the more predictive model in this research.

6.3.1 Statistical models

General linear regression and quantile models

General linear regression models (GLM) are common tools used in statistics in which it is aimed to predict the relationship between a dependent variable Y and a set of predictor variables X as specified in equation 6.1 below. In this equation, the betas are parameters which measure the individual effects of each X-variable when other variables are assumed constant and ε is an error term which is assumed to have a normal distribution with constant variance across different values of the X's. Hence, the GLM presumes a normal distribution where the mean medium and mode are the same, and the dependent variable Y as well as most X's are assumed to be continuous variables. An example of Y is Customer Value. General linear Models (GLM) have been used to predict CLV (Verhoef and Donkers, 2001).

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + . . . + \beta_P X_P + \varepsilon \qquad (6.1)$$

In circumstances when the distribution is not normal, then an alternative model should be considered, e.g. quantile regression (Koenker and Bassett (1978)). Quantile regression is a type of regression analysis, but instead of using the mean, the median (or other quantiles if required) is used. Due to CLV not always containing a normal distribution, Benoit and Poel (2009) used quantile regression for their research into predicting CLV.

Customer retention

Unlike customer value, which has a continuous target, customer retention is a binary target or dependent variable, which depicts whether the customer renewed their insurance or not. For this reason a logistic model was used such that if the logistic model dependent variable is denoted by Y and the vector of predictor variables by X, what is regressed is the odds-ratio of probability to renew or be retained P to the probability not to renew given by p/(1-p) so that the logistic regression equations are given by

$$\log\left(\frac{p}{1-p}\right) = b_0 + b_1 X_1 + \ldots + b_n X_n \Leftrightarrow p = \frac{e^{(b_0 + b_1 x_1 + \cdots b_n x_n)}}{1 + e^{(b_0 + b_1 x_1 + \cdots b_n x_n)}} \qquad (6.2)$$

Where b_0 is an intercept, $b_1, \ldots, b_n$ are the coefficients of the predictor variables $x_1, \ldots, x_n$ are observed values of the predictor variables X.

Data mining

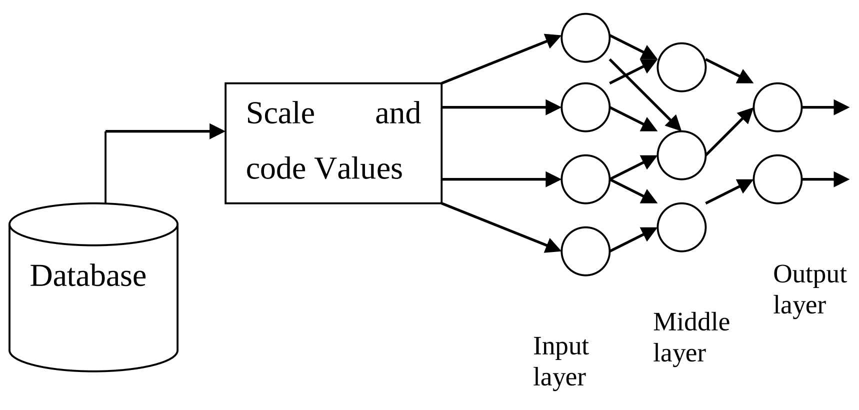
Two data mining techniques were involved in predicting customer value and customer retention:

- Neural networks
- Decision trees

> *Data mining is an interdisciplinary field that brings together techniques from machine learning, pattern recognition, statistics, databases and visualization to address the issue of information extraction from large databases.*
>
> *(France et al., 2002)*

Neural networks are data mining tools that work by training a network of neurons linked by connections to learn rules.

Figure 6.1: Diagram of a neural network



(Rygielski *et al*. 2002)

The most commonly used neural network is the multi-layer perceptron (MLP), which can be trained using the back-propagation learning rule (Han *et al.*, 2011). Neural networks have been used in different scenarios with large datasets, including CLV (Drew *et al*., 2001).

Some of the key concepts which determine a Neural Network (NN) are nodes, interconnections and architecture. A node is connection point which links input variables X's to other variables and ultimately the dependent variable Y. They are represented by the first, second and third columns of circles above. The complete design of a NN as shown in the above example is referred to as its architecture. A simple-perceptron NN is a connection of one column of nodes X's whereby each connection has a weight that can be varied as the NN uses information from an input dataset X to obtain desired values of Y. The multilayer perceptron includes the main input layer in the first column of nodes and a number of hidden or middle layers similar to the second column in the above example, which relate each variable to other variables within the set, and are such that the layers nearer to the final output layer are lumped-up variables, similar to principal component bundles of factors which account for most of the variation in the outcome variable Y. The output layer is typically a single node that captures the dependency of Y on the predictor variables. given these structures, the output Y is expressed as weighted sums of the X's as follows, respectively, for a simple and a multilayer NN

$$Y = \sum_{i=1}^{n} w_i X_i \text{ and } Y = \sum_{i=1}^{n} w_{ij} X_i$$

where w_{ij} are weights for the middle-level node connection from input *i* to contributing nodes *j*. NNs generally use back-propagation to learn the true values of the weights which produce the y-values to a reasonable degree of accuracy from the input dataset (X, Y). Back-propagation consists of calculating model estimation errors backwards from the outcome layer and continually adjusting the weights at each layer based on the equation

$$w_n = w_o + \lambda(d - y)x$$

where d is the desired outcome Y, y is the estimated outcome, e = (d-y) is the error, w_n is the new or updated weighting, w_o is the previous weighting, X = x is the vector of predictor variables going into the NN model and λ is a learning-rate parameter which is in the interval $0 < \lambda \leq 1$. In effect, the NN continually processes the data modelling algorithm at each interconnection by refining the weightings in order to minimize the errors to a negligible level when the learning process ends. This way, NNs avoid over-learning from the given data and approximate the Y-X relationship closely.

As well as neural networks, decision trees will also be used. A decision tree is another data mining technique where each branch node represents a choice between a number of alternatives, and each leaf node represents a classification or decision. Three different decision tree techniques were used

- Chi-square CHAID (*Chi-square Automatic Interaction Detection*)
- Gini Reduction - CART (*Classification And Regression Trees)*
- Entropy reduction - C5.0 algorithm

CHAID is one of the oldest tree classification methods, originally proposed by Kass (1980). CHAID can build trees with two more branches, based on a simple algorithm, where WSS is the total sum of squares after the split and TSS is the total sum of squares before the split.

$$WSS = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (y_{ij} - \bar{y}_j)^2 \qquad (6.3)$$

$$TSS = \sum_{j=1}^{g} \sum_{i=1}^{n_j} (y_{ij} - \bar{y})^2 \qquad (6.4)$$

where $\bar{y}_j$ is the mean value of y_{ij} in node j, for g groups and $\bar{y}$ is the overall mean.

The test or splitting criteria are the p-value of the F statistic for the difference in mean values between the g nodes generated by the split:

$$F = \frac{BSS/(g-1)}{WSS/(n-g)} \qquad (6.5)$$

where BSS=WSS-TSS

CART (Classification and Regression Trees)

Another popular decision tree technique is CART which was developed by Breiman, *et al*. (1984). CART uses Gini impurity, which measures how often a randomly chosen component would be labelled incorrectly. CART then applies a goodness of split criteria, at each split point and then evaluates the decrease in impurity (heterogeneity) caused by the split. At each given node, the probability distributions are:

$$\sum_{k} P_K \left(\sum_{all i, j i \neq j} p_{i/k} p_{j/k} \right) \qquad (6.6)$$

Where P_K is the proportion of the K^{th} leaf and $P_{i/k}$ is the proportion of observations in the category I (of the dependent variable) in leaf K

Entropy reduction - C5.0

The C5.0 models works by splitting the dataset based on which field provides the maximum information gain, the difference in the entropy of a node and the entropy after a node splits. The split criteria for these, is a measure of Entropy or gain ratio.
Entropy at a given node t:

$$Entropy(t) = -\sum_{j} p(j|t) \log(\mathrm{p}(j|t) \qquad (6.7)$$

Where p(j|t) is the frequency of class j at node t
Information gain:

$$GAIN_{SPLIT} = Entropy(p) - \left(\sum_{i=l}^{k} \frac{n_i}{n} Entropy(i)\right) \qquad (6.8)$$

$$SplitINFO = -\sum_{i=l}^{k} \frac{n_i}{n} log \frac{n_i}{n} \qquad (6.9)$$

Where parent node P with n records is split into k partitions; n_i is the number of records in partition node i.
The $GAIN_{SPLIT}$ measure the reduction in entropy achieved due to the split.

$$GainRATIO_{split} = \frac{GAIN_{Split}}{Split\ INFO} \qquad (6.10)$$

6.4 Measurement and data

6.4.1 Data description

The data used comes from an established UK car insurance company between January-May 2010 (inclusive). The data set contained 189,798 rows of data from policies that were due for renewal. Customers that had cancelled their insurance mid-term were not considered in this analysis. The data provided from the insurance contained the explanatory variables in Table 6.1 below.

Smith *et al* (2000) used the premium difference between the renewal premium and old premium. This is not available for targeting customers mid-term, but a customer value metric can be calculated.

As price comparison sites are a new phenomenon, a graph will be produced to measure the impact of the marketing communication (media source) on renewal rates. Apart from online and aggregators (which are computer generated), the media source is recorded by the company when the customer answers a specific question during their quote process i.e. 'How did you hear of the company?' The marketing campaign choices are as follows:

Table 6.1: Explanatory variables

Variable	**Category**
Age group	<=23, 24-28, 29-32, 33-37, 38-46, 47+
Allowed to contact policy holder	Yes or No
Car age	1-2, 3-5, 6-9, 10 +
Car colour	Blue, black, brown, green, grey, pink, red, silver, white, yellow, other
Claimed on insurance	Yes or No
Gender	Male, Female, other
Married	Single, married, cohabiting
Marketing cost	£'s (continuous)
Marketing Source	TV, Directory, Print, radio, win back, door drop, aggregator, online, personal referral
No Claims Bonus (years)	1, 2-3, 4-5, 6-8, 9-10, 11+
No Claims Bonus Protected	Protected or No

Number of drivers	1, 2, 3 and 4+
Pay method	Direct Debit, Credit Card, in Full
Renewal year	1, 2, 3 4+
Social grouping	1-wealthy, 2-prosperity, 3-comfortably off, 4-moderate means, 5-Hard pressed, 6-other
Total claims costs	In £'s (continuous)
Total premium	In £'s (continuous)
Type of insurance cover	Comprehensive / Third party fire and theft/ third party only
UK region	East Anglia, East Midlands, Greater London, N Ireland, North, North West, Other, Scotland, South West, Wales, West Midlands, Yorkshire & Humber, Other Southeast
Value	In £'s (continuous) *(equation 6.11)*
Vehicle group	1-4, 5-6, 7-12, 13-20, 21+

Table 6.2: Marketing source

Media source	**Description**
AGGREGATOR	Price comparison site
DIRECTORIES	Telephone Directory
ONLINE	Web based search and banner ads
OUTDOR	Ambient (e.g. Poster, back of bus)
PRESS & MAGS	Newspapers and magazines
RADIO	Radio
REFERRAL	Recommendations from a friend/family
TV	Television advertising
UNKNOWN	Unknown
WIN BACK	Ex-customers who have returned

These choices gives some insight into customer channel behaviour, as online and aggregator are internet based channels, with aggregator being the intermediary, and the other sources representing a mixture of internet and telephone channels.

6.4.2 Targeting valuable customers

The segmentation strategy will apply two different scenarios, RFM and a three dimensional strategy based on actual value, CLV and predicted renewal. The RFM strategy will need to be amended as 'R' will not be relevant as all customers would be on an annual contract.

The CLV strategy would be extending Verhoef and Donkers (2001)'s two-dimensional approach of customer segmentation (potential and customer value) to include predictive retention rate (Kim *et al*., 2006; Hwang *et al*., 2004 and Guelman *et al*., 2012). Using predictive figures as well as actual figures provides a powerful tool in customer analysis. Predictive statistics figures are never 100% accurate, so using actual customer value would highlight customers that have a negative equity. This analysis will use a mixture of data mining and statistical techniques to predict the probability of the customer retaining their car insurance policy as well as calculating their customer lifetime value. This information will enable policy decision analysis and help marketers with their customer targeting.

Customer value

Cross-selling products is not an option for the company, as they focus on car insurance only, whereas ancillaries purchased are not considered important enough by the company for the value metric to be used.

To calculate the value of a customer within an insurance company, claim costs as well as marketing costs need to be considered. Claim costs can be calculated as:

$$\begin{matrix} claims \\ costs \end{matrix} = \begin{matrix} Total \\ Paid \end{matrix} - \begin{matrix} Total \\ Recovered \end{matrix} + \begin{matrix} Payment \\ reserve \end{matrix} - \begin{matrix} Recovery \\ Reserve \end{matrix} \qquad (6.11)$$

Total paid refers to any outgoings incurred from the policyholder, either through an accident or a car being stolen. The total recovered would be the amount recovered from the claim, typically from the third party's insurance company where the car accident was not the policyholder's fault. If a claim has not been settled, then payment and recovery reserves will have values that need to be considered. Staff employed by the insurance company, initially calculate the claim pay out reserves and recovery reserves. With the claims costs considered, customer value will be calculated as:

$$\begin{matrix} Customer \\ Value \end{matrix} = \begin{matrix} Total \\ Premium \end{matrix} - \begin{matrix} Acquisition \\ marketing\ costs \end{matrix} - \begin{matrix} Claim \\ Costs \end{matrix} - \begin{matrix} Retention \\ marketing\ costs \end{matrix} \qquad (6.12)$$

Total premium will be the total premium gathered. Acquisition marketing costs are calculated on a direct basis, depending on the media choice the customer quoted when purchasing. Claim costs will be the total cost of the claim (*equation 6.11*). Retention marketing costs cover any marketing costs associated with the customer to maintain their custom. The value metric can be applied to each individual customer, which can be used to predict customer lifetime value. As retention marketing has not been implemented, these costs would be £0.

Variable selection for value model

To determine which variables should be used to predict customer value from table 6.1 F-tests were carried out. F-tests are significance tests used to determine whether the variable contributes to the accuracy of the model.

$$F = \frac{\frac{\sum_i n_i(\overline{Y_\iota} - \bar{Y})^2}{(K-1)}}{\frac{\sum_{ij}(Y_{ij} - \overline{Y_\iota})^2}{(N-K)}} \tag{6.13}$$

Where $\overline{Y_\iota}$ is the sample mean in the i^{th} group, n_i is the number of observations, $\overline{Y}$ is the overall mean, k denotes the number of groups, Y_{ij} is the j^{th} in the i^{th}out of K groups and N is the overall sample size. The final variables were applied to all of the value variables in Table 6.1 to obtain the list of significant variables in Table 6.3 below.

Winsorisation

As well as different statistical techniques, this study will also consider Winsorization. Winsorization is a robust regression technique designed to control the influence of outliers. Values greater than the 98th percentile are set equal to the 98th percentile, therefore the top 2% become £3903 and the bottom 2% become -£2239. The 98th percentile was chosen as this made commercial sense to remove any extreme negative values, while using as much actual data as possible.

For the validation tests, the data was split into two (using a simple random technique), 70% for training and 30% for validation. The explanatory variables were then used for general linear regression, CHAID analysis and neural networks. Due to there not being an option to use different decision tree techniques for a continuous target variable, both neural networks and decision trees will use the default settings. Decision trees will use F-test with a significance level 0.2 whereas for neural networks, 'this is a multiplayer perceptron with no direct connections and the hidden layer is data dependent' (SAS, 2003, p.67).

Comparing their statistical errors and residuals, the most predictive technique will be found to predict customer value. Statistical errors are the amount the observed error differs from its expected value, whereas the residuals are the deviations of the dependent variable observations from the fitted function.

Value performance metrics

Quantile regression uses the median, so using the mean absolute error (MAE) and root mean square error (RMSE) statistics to measure model performance, as used by Verhoef *et al*. (2007) and Leeflang (2000), is not feasible in this instance. Therefore to measure the accuracy of the predictive value against the actual value, the hit-rate criterion proposed by Donkers (2007) will be used.

The hit rate criterion categorizes all of the customers, based upon their actual value, into four equal segments with decreasing level value. These segments are then used to compute the percentage of predicted values which fall into the same category as their actual value (top quarter = £1373 +, second quarter £1372 to £665, third £664 to £322 and bottom quarter = £321 and less).

Retention analysis

Data specification

To predict the probability of a customer renewing their insurance with the company, three data mining tools were used: neural networks, decision trees and logistic regression. As value was one of the explanatory variables for retention, there would be cause for outliers affecting its distribution (e.g. customers having a claim for over £100,000). To reduce the influence of outliers a 2% Winsorization was applied, as before i.e. values greater than the 98th percentile are set equal to the 98th percentile. The data was split into two, 70% for development and 30% for validation. This means that 70% of the data will be used to build the model, with 30% of the data to test for stability.

To test for stability and accuracy for the different data mining techniques the root average squared error (ASE) will be used. The root ASE represents the standard deviation of the differences between the predicted and observed values.

$$Root\ ASE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \qquad (6.14)$$

where X_{obs} is observed values and X_{model} are modelled values.

Retention model variable selection

The next stage involved using a variable selection node to pick out the variables that have a significant impact on whether a customer renews or not (Table 6.5), which entailed using chi-square (equation 6.5).

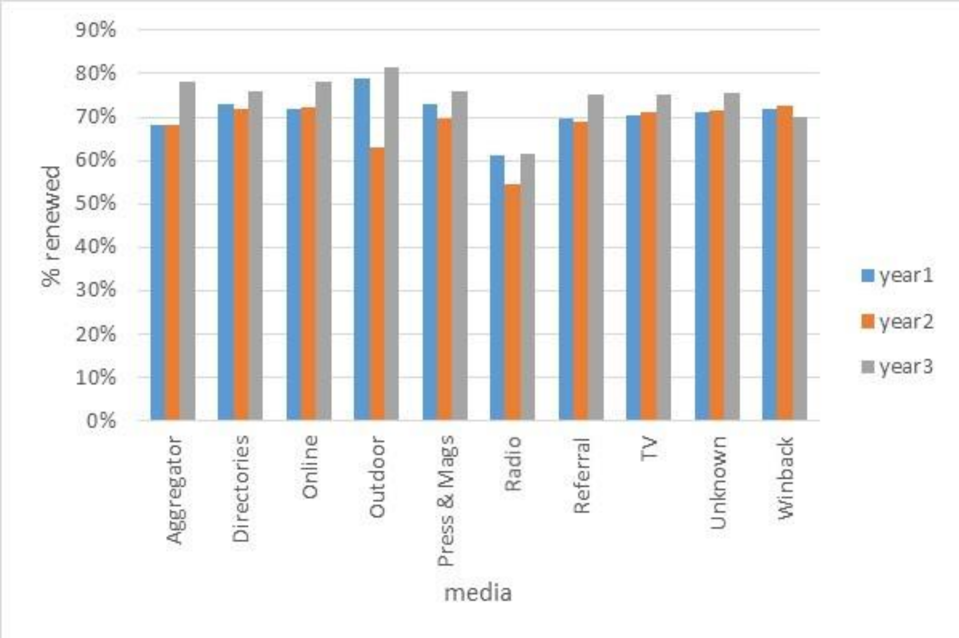
Modelling techniques

For the data mining exploration of the data, an off the shelf application was used, SAS Enterprise Miner, to ensure the application of neural networks and decision trees on the data set. The data mining techniques will be used to compare against logistic regression. For neural networks a multilayer perceptron (MLP) will be used. MLP is a feed forward artificial neural network, which consists of multiple layers of nodes, each one connected to the next. The research will use three different neural network techniques: Levenberg-Marquardt; Conjugate gradient, and Quasi-Newton). So that the most predictive decision tree technique is used for comparison with neural networks and logistic regression, three decision tree techniques will be compared against each other: Gini; Entropy and, Chi-square. These different decision tree techniques were compared with each other, respectively. Lift charts were produced to discover which decision tree and neural network technique produced the strongest discriminatory power between those who renewed and those who did not. Finally, the explanatory variables chosen from the variable selection process were used for logistic regression, decision tree analysis and neural networks, based on the likelihood of a customer renewing their car insurance policy.

6.5 Results

6.5.1 Renewal rates by media channel

Figure 6.2: Retention rates by marketing communication across various channels



From Figure 6.2, customers gathered via aggregator sites during their first renewal cycle (rnyear 0), are less likely to renew than those gathered through other acquisition channels (excluding radio). This trend does not carry on in subsequent years as the aggregator channel performs more favourably compared to the other acquisition channels. This differs from Verhoef and Donkers' (2005) research, which found that the effects do not carry on over a number of years.

6.5.2 Customer value results

Figure 6.3: Customer lifetime value distribution

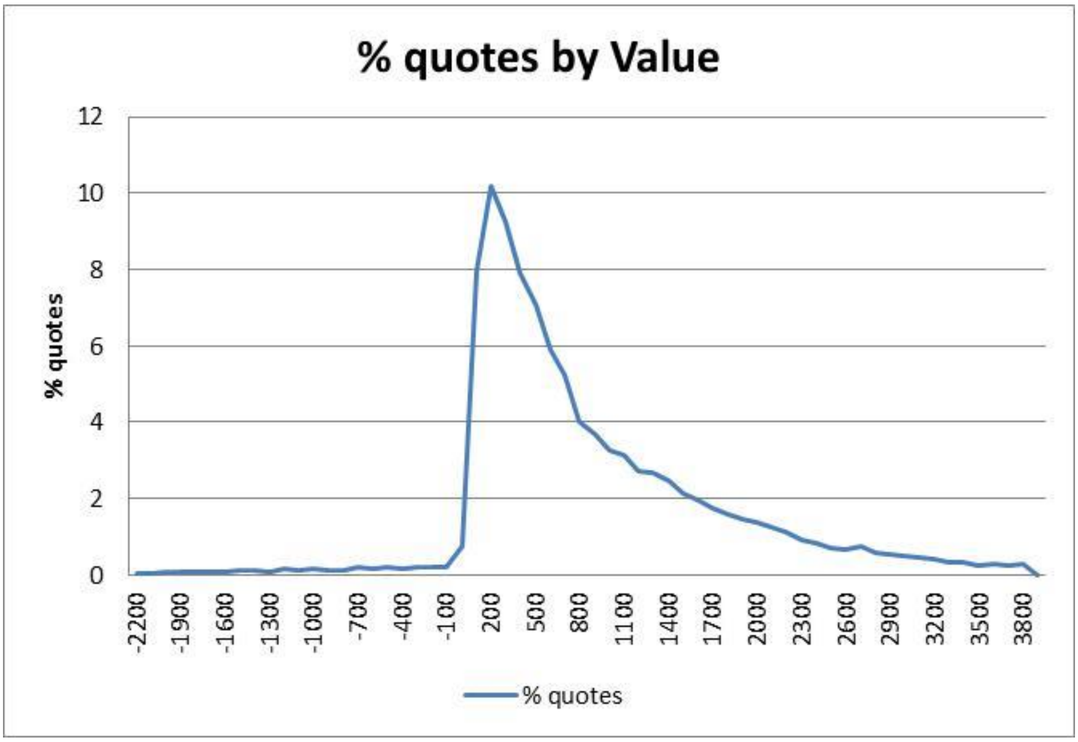


Figure 6.3 shows that CLV does not exhibit a perfect normal distribution. The distribution tends to be positively skewed, so the linear regression model and the quantile regression model should produce different statistical models. Descriptive statistics of actual customer value are displayed in table 6.4.

From figure 6.3, there is a peak at between the £200 and £300 value mark. This peak can be attributed to the company's average value for customers who did not renew. These people have most probably not claimed, so have generated a positive value for the company. For this peak to move to a higher positive figure, more customers will need to renew.

The median customer value on the full population is £661, whereas the mean is £790. When applying Winsorization, using the lower and upper 2 % (-£2239 and £3903), the median stays the same, £661, but the mean changes to £907. The change in mean demonstrates that outliers on the negative side affect the mean more than those on the positive side. This confirms the view in section 6.2 regarding high value claim costs.

Variables used for the neural networks, GLM and quantile models

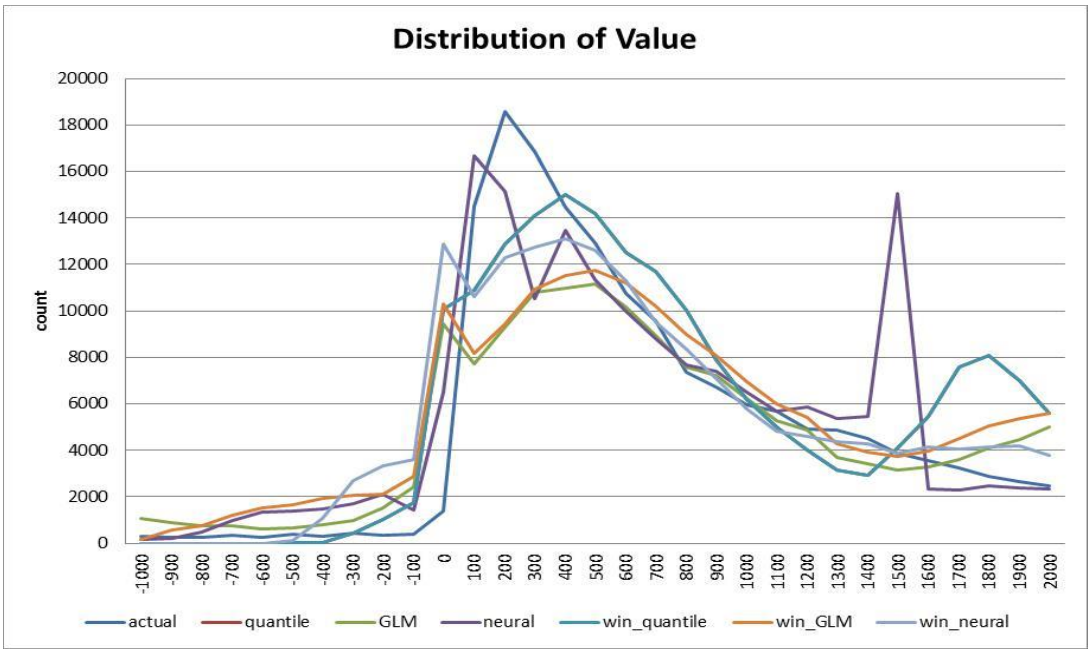
For the value model, each variable from table 6.1, F-test criteria was used.

Table 6.3: Variables used for value models

Variable	DF	Type III SS	Mean Square	F Value	Pr > F
Age group	5	4769857792	953971558	33.94	<.0001
Car age	3	1281991382	427330461	15.2	<.0001
Claimed on insurance	1	83074922897	83074922897	2955.55	<.0001
Type of insurance cover	2	297159141	148579571	5.29	0.0051
Married	3	356830564	118943521	4.23	0.0053
Marketing Source	10	733522463	73352246	2.61	0.0036
No Claims Bonus	5	4190869720	838173944	29.82	<.0001
No Claims Bonus Protected	1	175404478	175404478	6.24	0.0125
Number of drivers	4	712576459	178144115	6.34	<.0001
Pay method	2	1187998704	593999352	21.13	<.0001
UK region	12	2880284164	240023680	8.54	<.0001
Renewal year	3	47172816759	15724272253	559.42	<.0001
Gender	1	484498819	484498819	17.24	<.0001
Vehicle group	4	3386789271	846697318	30.12	<.0001

With the variables created the quantile model (Appendix 6.1), the GLM (Appendix 6.2) and the Winsorised GLM (Appendix 6.3) were created.

Figure 6.4: Distribution of the predicted customer value



Due to lack of variance in decision trees, the plot for these readings had to be omitted.

Table 6.4: Descriptive statistics of modelled value

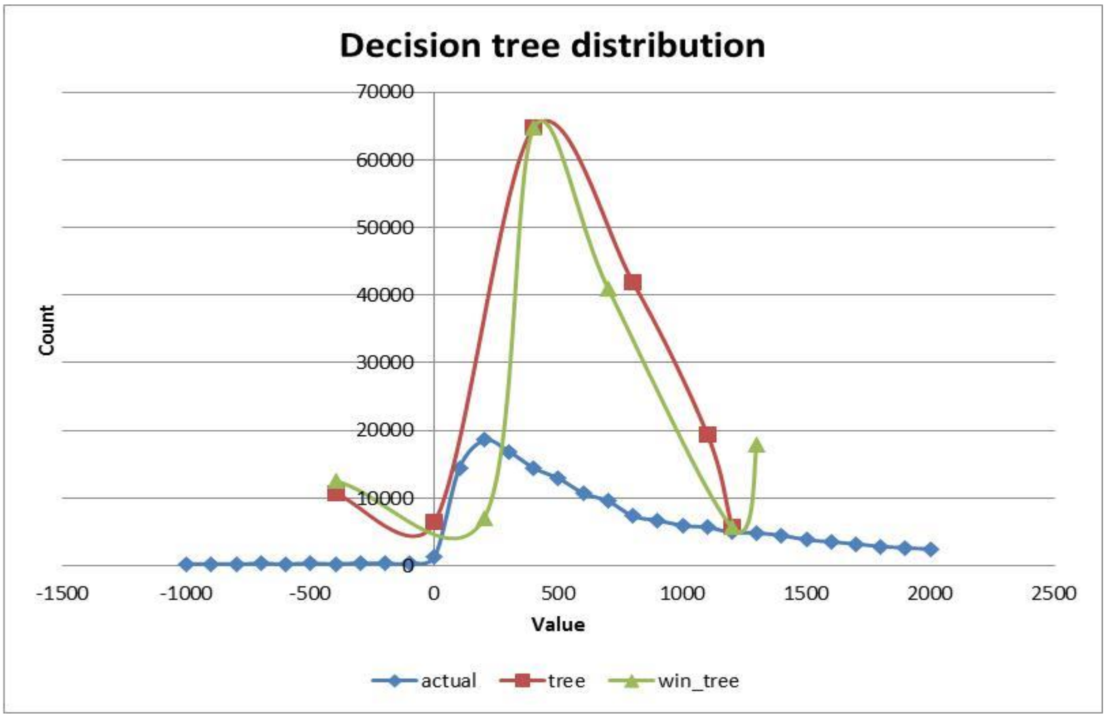
Value	mean	Std. dev.	max	min	skewness	kurtosis
Actual	£786	£2595	£11,683	-£178,427	-26.7	1431.3
GLM	£786	£1059	£3,576	-£2,614	-0.36	0.164
Quantile	£894	£664	£3,120	-£549	0.56	-0.80
Neural	£867	£784	£4,110	-£1,794	0.67	0.39
Tree	£912	£727	£2,100	-£437	0.41	-0.62
Actual (win)	£908	£1047	£3,903	-£2239	0.377	1.99
GLM (win)	£909	£807	£3,388	-£1,296	0.30	-0.54
Quantile (win)	£894	£664	£3,120	-£549	0.56	-0.80
Neural (win)	£910	£817	£4,575	-£562	0.82	0.07
Tree (win)	£913	£742	£2,097	-£480	0.34	-0.69

From Table 6.4, it can be observed how the extreme values of claims costs attributed to the company and that each different technique has a different mean. When considering actual value, it can easily be observed.

From Figure 6.4, neural networks exhibit a high number of observations at £1500, which may contribute to the predictive accuracy model. This peak is removed when neural networks is applied to the Winsorised dataset. Also demonstrated in the above chart is the effect of applying different statistical techniques to Winsorised and non-Winsorised data sets.

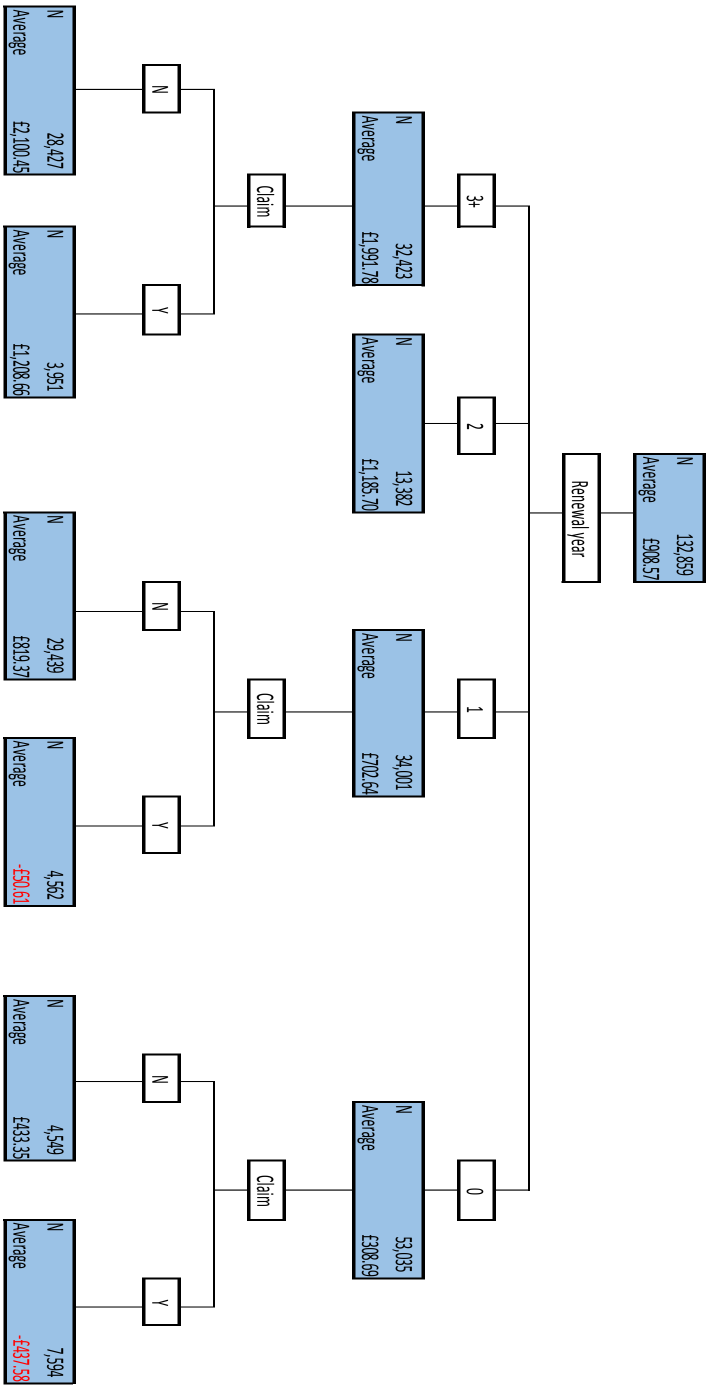
- GLM model on the Winsorised data set tends to follow the actual customer population than the non-winsorised GLM
- Winsorization has no affect when building a quantile model and graph 6.4 confirms this
- For neural networks, Winsorisation provides a more stable distribution and removes the peak at £1500

Figure 6.5: Comparison of decision trees



The lack of values from the decision trees does limit its overall predictive power.

Figure 6.6: Decision tree analysis of value



Decision trees use the most powerful explanatory variable that will predict the target variable, in this case customer value, to create its initial branch. From the value decision tree (Figure 6.6), it can be shown that renewal year is a very strong predictor for value, which is expected as the longer the customer stays with the company, the more valuable they are. Also as expected, claim is the second strongest predictor, as customer value would be decreased if the company had to pay money out. From the decision tree (figure 6.6) it can be shown that customers that have stayed with the company for 3 or more years and have made an insurance claim, on average, are still worth more than people who have been with the company for less time regardless of whether they have made a claim or not. This demonstrates a relationship between customer retention and customer value within the car insurance industry

Hit rate results

Figure 6.7: Overall hit rate

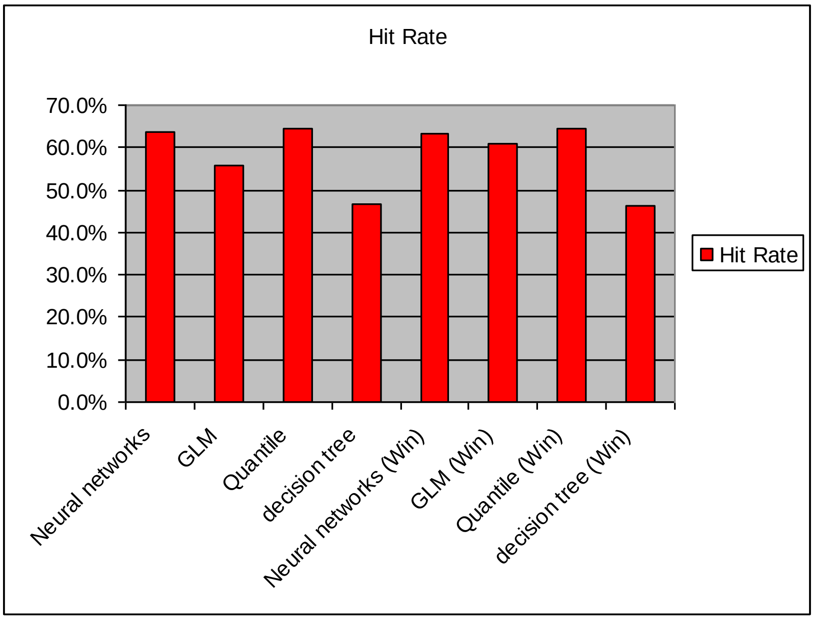


Table 6.5: Hit rate results

Model	**Segment 1 (£1373+)**	**Segment2 (£664-£1373)**	**Segment 3 (£321-£664)**	**Segment4 (under £321)**	**Overall Hit Rate**
Neural networks	*79.0%*	*54.4%*	*35.7%*	*85.0%*	*63.5%*
GLM	*79.2%*	*42.4%*	*25.9%*	*76.4%*	*55.9%*
Quantile	*77.7%*	*57.6%*	*43.6%*	*79.8%*	*64.7%*
Decision tree	*63.8%*	*52.8%*	*45.6%*	*24.6%*	*46.7%*
Neural networks (Win)	*79.2%*	*50.8%*	*37.4%*	*85.6%*	*63.2*
GLM (Win)	*81.3%*	*52.0%*	*32.1%*	*78.5%*	*60.9%*
Quantile (Win)	*77.7%*	*57.6%*	*43.6%*	*79.8%*	*64.7%*
Decision tree (Win)	*79.3%*	*33.9%*	*45.6%*	*26.5%*	*46.3%*

The hit rate criterion categorizes all of the customers, based upon their actual value, into four equal segments with decreasing level value. These segments are then used to compute the percentage of predicted values which fall into the same category as their actual value e.g. for neural networks in the first segment, 79% of the highest predicted values belong in the top segment. Figure 6.7 and table 6.4 both confirm that quantile regression performs the strongest, overall, when calculating CLV, which corresponds to the research conducted by Benoit and Poel (2009).

From table 6.5, it can be shown that the model loses most of its power in the third segment for all models. The differences can be partially explained by how each model treats its target variable, the mean for GLM and decision trees and the median for the quantile regression and the different techniques applied, especially for neural networks which uses a ‘black-box’ approach. Also demonstrated is the fact that Winsorized data models perform better at the extremes when considering hit rate analysis. Finally, Winsorized data only seems to improve the general linear model when considering the overall effectiveness of the models. The top segment tends to differ from Malthouse and Blattberg (2005), who found that out of the top 20% of the customers 55% would be misclassified.

6.5.3 Customer retention

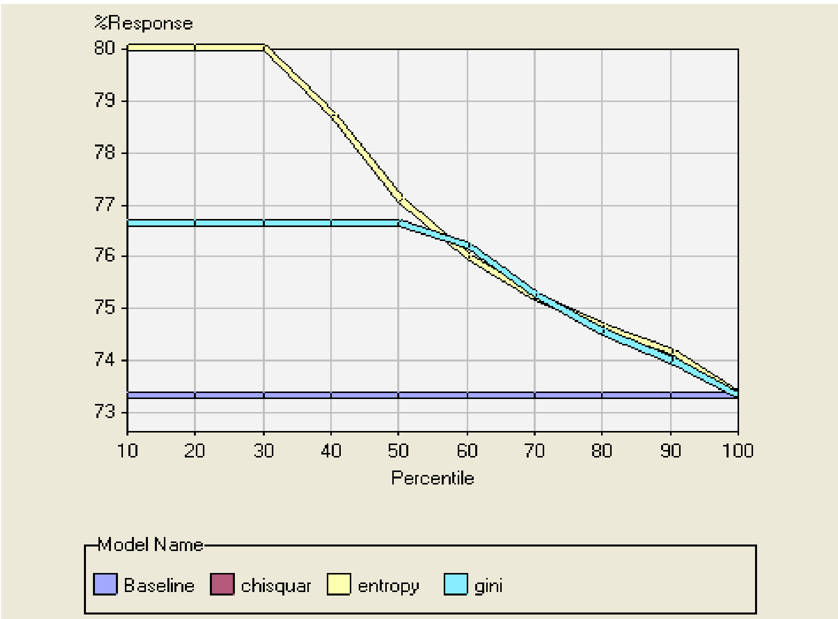
Decision tree results for retention

Three different decision tree techniques (Gini, Entropy and Chi-square) were tested to find which provided the most powerful predictive technique.

Table 6.6: Decision trees standard errors

Tree	Root ASE	Test Root ASE	Misclassification Rate	Leaves
Gini	0.4404	0.4413	0.2667	5
Entropy	0.4390	0.4398	0.2667	7
Chi square	0.4404	0.4413	0.2669	4

Figure 6.8: Comparison of the three different decision tree models for customer retention analysis



From Figure 6.8, all three different decision tree techniques outperform random sampling (baseline). Both the Gini method and the Chi square methods produce the same plot, so chi-square is hidden behind the Gini plot. By targeting the top 30% of the customers we would expect:

- 80% renewal rate for Entropy method
- 76.7 % for Chi-square and Gini
- 73.3% Baseline (not targeting)

From the above statistics, and using table 6.6, the Entropy method is to be applied for comparison against neural networks and logistic regression as it has the highest 30% targeting figure and lowest root ASE. Also, from table 6.6 it can be observed that the root ASE and the test ASE are similar, which shows stability between the build data set and the validation data set (test).

Logistic regression model results

Using chi-square, the explanatory variables (from table 6.1) were reduced into the ones that were most predictive.

Table 6.7: Variables used for retention

Variable	Degrees of freedom	Wald Chi-square	Pr > ChiSq
Age group	5	160.177	<.0001
Car age	3	491.2033	<.0001
Claimed on insurance	1	382.426	<.0001
Social grouping	18	859.9115	<.0001
Marketing Source	10	148.835	<.0001
No Claims Bonus	5	352.0623	<.0001
No Claims Bonus protected	1	24.9487	<.0001
Number of drivers	4	244.1347	<.0001
Allowed to contact policy holder	1	415.6187	<.0001
Pay method	2	143.4986	<.0001
UK region	12	702.3182	<.0001
Renewal year	3	799.0322	<.0001
Gender	1	372.2064	<.0001
Value	1	42.9459	<.0001
Vehicle group	4	38.5353	<.0001

With the explanatory variables chosen a logistic model was created (Appendix 6.4).

Neural networks results for retention

Using the variables from table 6.7, three different neural network methods were used to predict customer retention:

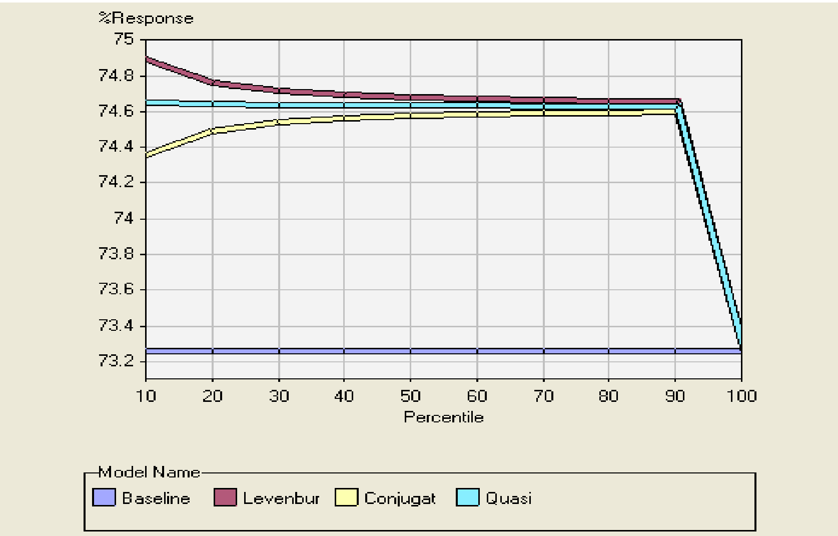
- Levenberg-Marquardt
- Conjugate gradient
- Quasi-Newton

The following statistics were produced:

Table 6.8: Neural networks statistics

Network	Root ASE	Test root ASE	Misclassification Rate	Schwarz Bayesian Criterion
Levenburg	0.44568	0.44735	0.26529	157878
Conjugate	0.44593	0.44701	0.26548	158014
Quasi	0.44578	0.44730	0.26508	157931

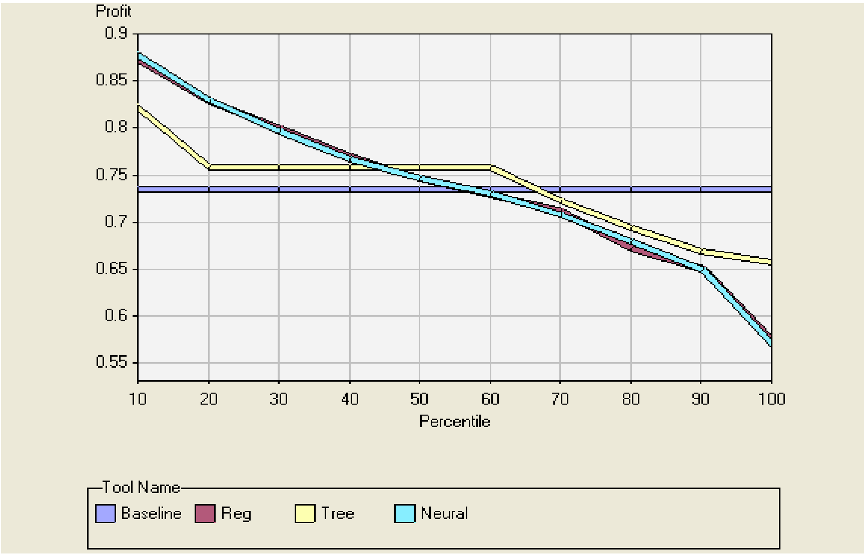
Figure 6.9: Comparison of the three different neural network models for customer retention analysis



From Figure 6.9, all three techniques outperform random (baseline), with the Levenberg-Marquardt outperforming the Conjugate gradient and Quasi-Newton techniques. From table 6.8, the Levenberg-Marquardt technique also has the lowest root ASE compared to the other two techniques. Also, from table 6.8, it can be shown the model is stable between the build and validation data sets, root ASE (0.44568) and the test root ASE (0.44735). Levenberg-Marquardt will therefore be the technique used to compare decision trees and logistic regression.

Comparison of neural network, decision trees and logistic regression

Figure 6.10: Non-cumulative profit



With a high retention rate for the base line, targeting can still provide improved results. All three techniques perform better than random, with neural networks and logistic regression outperforming decision trees. Both neural networks and logistic regression perform well and demonstrate that targeting the top 50% percentile would give a better performance than choosing a random sample.

Figure 6.11: Comparison of neural network logistic regression and decision tree models for customer retention analysis (non-cumulative)

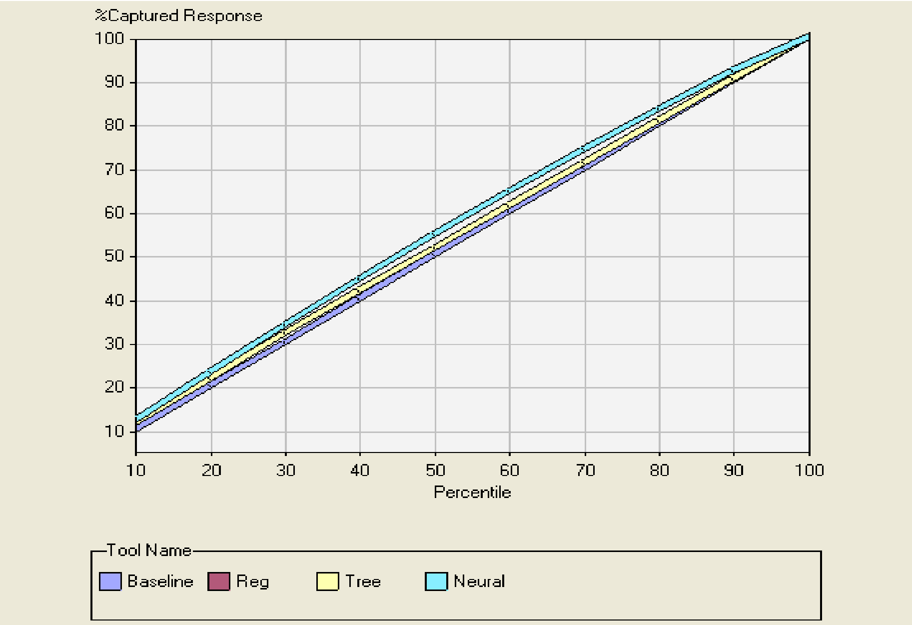


Table 6.9: Results of the three different techniques

Tool	**Root ASE**	**Test Root ASE**
Neural Network	664.776	671.601
Trees	746.709	751.497
Regression	656.614	665.936

Figures 6.10 and 6.11 demonstrate the comparative predictive powers of the different data mining techniques. From the lift chart (Figure 6.11) above, all three techniques can produce an improvement rather than just sampling a random set of customers (baseline). By targeting the top 20% percentile, the identification rate of people who are likely to renew is as follows:

- Regression 23.15%
- Tree 21.76%
- Neural networks 23.22%
- Random (baseline) 20%

From figure 6.11, the further down the percentiles, the more customer renewals are found, until the 100% mark is reached. Locating customers likely to renew compared to customers less likely to renew will influence how a company treats their customers and could help them market more favourable offers.

Table 6.9 shows that a general liner model has the lowest root ASE, thus outperforming neural networks and decision trees, and would therefore be the best predictive technique for value. Figure 6.11 also shows that the decision tree performs adequately, which as Donkers et al (2007, pp.182) note, ‘will provide a comforting idea to practitioners who use relatively simple models'. Also, from table 6.9, the root ASE and the test root ASE do not differ too much, which demonstrates stability in the models.

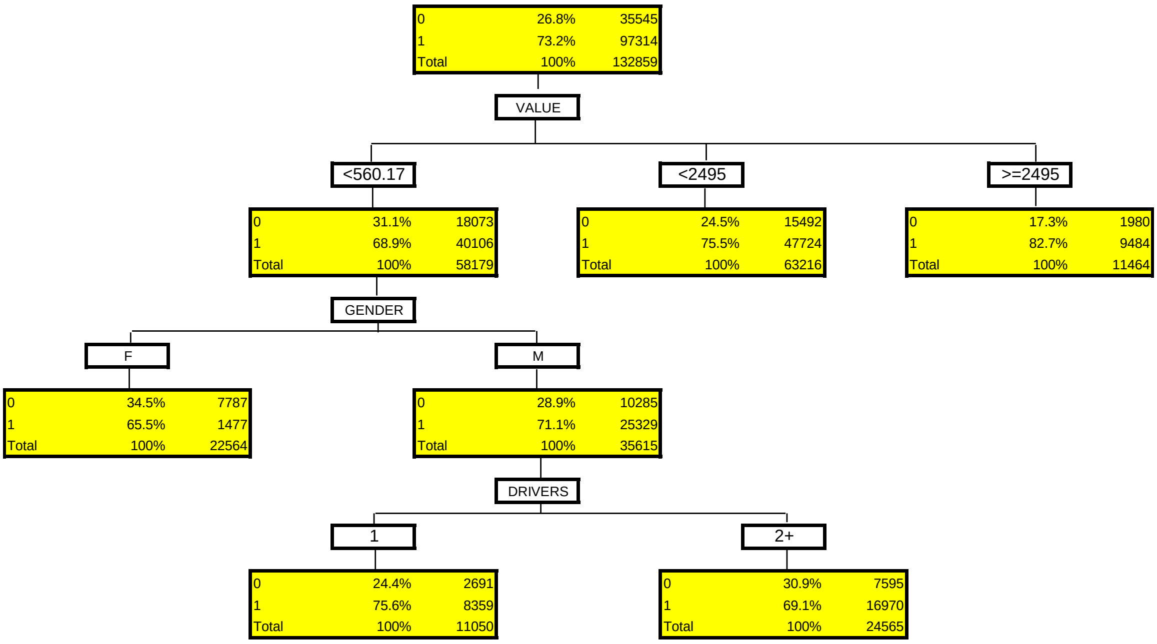


Figure 6.12: Decision tree analysis of retention

From figure 6.12 it can be shown that value is a very strong predictor for renewal. From the decision tree we note that:

- Customers whose value is greater than £2,449.50 are more to renew than any other group (82.7%).
- Customers whose value is between £560.17 and £2499.49 are as likely to renew as male drivers with one driver on the policy whose value is less than £560.17.

These customers are a slight improvement on picking out a random sample.

The customers least likely to renew are females whose value is less than £560.17 and male drivers whose value is less than £560.17, and with two or more drivers on their policy. In other words, the more money the company has made from the customer, the more likely they are going to renew.

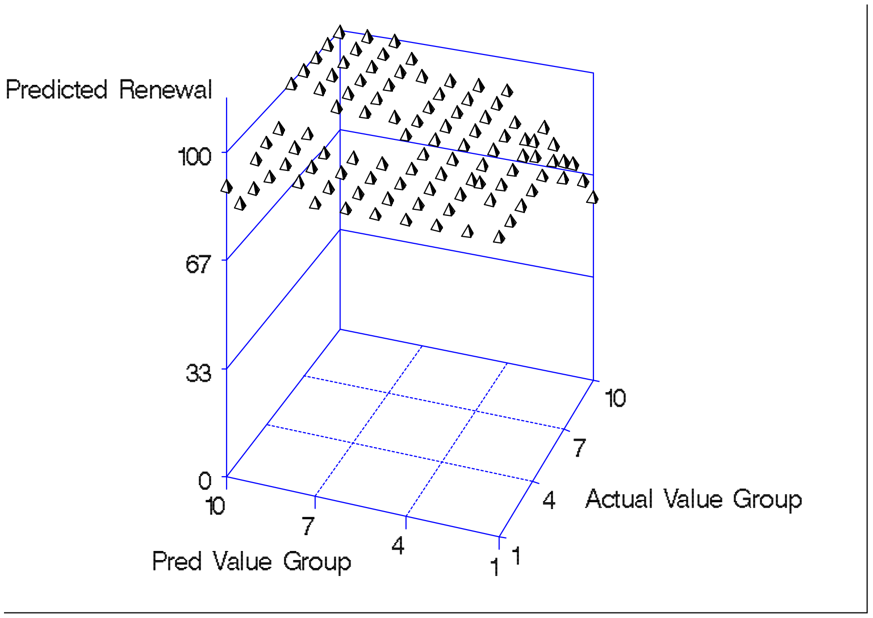
6.5.4 Customer segmentation

Businesses are aware of the 80/20 rule, where 80% of profits are generated by the top 20% profitable customers and 80% of costs are produced by the top 20% of unprofitable customers (Duboff, 1992). This awareness can lead to segmenting such customers where possible. Within the claims department, using actual customer value can highlight the highly unprofitable customers.

The decision tree analysis (Figures 6.6 & 6.12) highlights the power of RFM segmentation. The strongest predictive variable for customer value (figure 6.6) was renewal year, and for retention (Figure 6.12) was customer value. As frequency within the insurance industry is on an annual basis, this leaves R (renewal year) and M (actual customer value).

The second approach uses the actual value, the most accurate predictive value model (quantile), and the strongest customer renewal model (neural networks).

Figure 6.13: Three dimensional segmentation plot



Key: - Value groups were split into deciles based on value, each with the same number of customers in each group, with group 1 being lowest value and group 10 the highest. Predicted renewal is grouped into deciles, so 100 equal 90-100% chance of renewing.

From figure 6.13 it can be observed that all of the customers have a high renewal rate, with the customers in the highest predicted and actual value groups (10), being the most likely to renew. Also it can be seen that as the customers' predicted and actual value decrease, the less likely they will renew. Using actual value also demonstrates the limitations of using just predicted value.

Figure 6.13 is able to provide further insight into planning strategies. It can be used so that those customers with a high probability to renew do not need to be contacted. Another group not to contact would be those with a low actual value. A simple group to target could be those with a mid-range of likelihood to renew with a high predicted value and an above than average actual value.

6.6 Summary and conclusion

The main purpose of this chapter is to increase the understanding of the effect of aggregators within the UK car insurance industry business model (objective 2). The different scenarios reposition the case company with regards to its future growth and profitability.

Firstly, this chapter was created to fill in knowledge gaps within the UK car insurance industry, within a price comparison site environment, regarding retention strategies. Within insurance, customer renewal strategies need to focus on renewing low risk customers, rather than all customers, which differs from Reicheld (1996)'s general view. This chapter highlighted that customers behave differently depending on where the customers responded to marketing with price comparison sites to be considered as just another media type. Customers from aggregators tend to have a strong renewal rate, which demonstrates that companies should implement or continue their customer renewal strategy.

The starting point of any CRM project within the car insurance industry, is to locate which customers will generate the most profit to the company. Figure 6.3 demonstrates that a low risk customer who renews their insurance, may not initially seem to be the most profitable, but over time, may end up being very profitable for the company.

As shown in Figure 6.6, customers that have been with the company three years or more are more valuable than customers in their first year, even if the customers in the three-plus years segment have claimed. This statistic should not be ignored by the company and thus highlights the importance of renewing the 'right' customer. With the two different segmentation techniques outlined here, it is possible to propose strategies for enhancing customer profitability. The simplicity of using decision trees to create an amended RFM strategy (RM) provides a clear simple strategy for implementation when dealing with the customers the company would most like to evolve into high value customers.

CLV has been used on its own for customer segmentation, however as highlighted by the hit rate analysis there is some uncertainty in the results. With budget constraints in place in most businesses, the ideal would be to create more focussed segments that incorporate actual customer value and predicted renewal rate for a more targeted campaign.

Also, this chapter also focussed on creating a segmentation based on actual customer value with the modelling of customer predicted renewal rates and customer potential value. The chapter compared different statistical methods along with data mining to produce the most accurate modelling technique. With respect to predicting customer value, quantile regression analysis appears to have the greatest predictive capability. This chapter used Winsorisation to remove the long tailed distribution shape, and although this technique does improve the models using the mean, the quantile regression still outperformed these. For predicting renewal rates, neural networks outperformed decision trees and logistic regression, albeit slightly.

One of the main uses of CLV is for segmentation purposes and this chapter believes that just focussing on value neglects another key metric: whether someone is likely to purchase from the company again. Considering this extra dimension will give the decision maker more information for a better return of investment when focussing on their renewal strategy. This extra dimension may prompt the question 'Is it worth contacting customers that have an over 85% chance of renewing the insurance again?' and so possibly lead the company to focus on the people in the middle range.

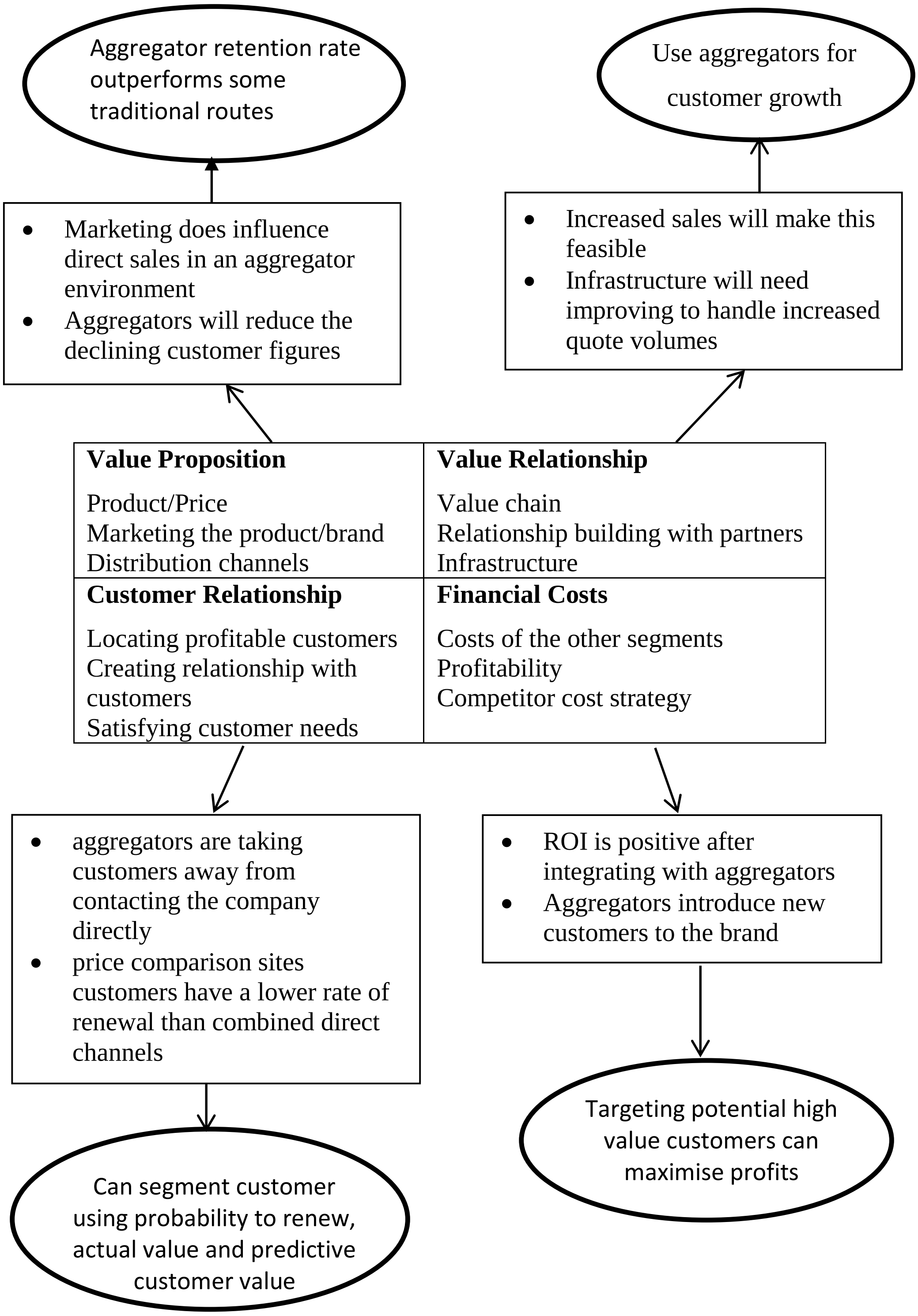
Finally, this chapter suggests that a combination of statistics and data mining are appropriate depending on the strategy. For a company using an RFM technique, the segmentation suggested from decision trees may be ideal. If a company would like to score up all customers and have a more granular approach then neural networks would be used for predicting customer renewal rate, with quantile regression for predicting customer value. If the marketing team is willing to use data mining and statistical techniques, then a more granular strategy could be implemented. If a customer has a high potential value, then that customer could be involved in a customer retention strategy of a deal with a free ancillary product.

The simplicity of the RM model is based on statistical analysis, which may give some comfort to managers proposing a strategy.

Implications of the results for car insurance modelling

The above results inform the development of an Integrated Business Re-intermediation Model (IBRM) for car insurance as shown in Figure 6.14 below, where the rectangular boxes denote the insights from the results in Chapter 5 and the elliptical boxes denote those from this chapter.

Figure 6.14: Developing the IBRM (part 2)



6.7 Limitations and Further Research directions

This study has limitations that future studies could focus on. Firstly, the results of this study are focussed on the UK car insurance, so other industries have not been accounted for. Secondly, there have been minimal discussions of customers' reasons for non-renewal with the company. This information could provide valuable insight and enhance the models which future research may attempt to incorporate into their strategy. Finally, with price comparison sites still new, their full impact on CLV and retention models could not be fully explored. Further research could study whether customers who come through aggregators can be treated with the same retention program as customers who approach directly, or whether a different strategy would be more appropriate.

Chapter 7: Car insurance marketing in the price comparison environment

7.1. Introduction

Financial price comparisons sites (aggregators) have changed the UK car insurance industry dramatically. Whether a car insurance company decides to join an aggregator or not, their marketing strategy has to adapt to this new environment. This chapter reviews different marketing techniques, to develop a new marketing framework for the UK car insurance industry. An application within the car insurance industry found that personal referrals affect aggregator acquisition rates, whereas web marketing and DM marketing mainly affect direct acquisitions. The research also notes that ambient marketing affects aggregators and direct channel acquisitions as well as customer renewal rates. The discoveries have strong implications for the UK car insurance industry by way of focussing their marketing strategies by using the framework proposed.

Marketing provides a pivotal role in business, and the arrival of aggregators has meant strategic marketing is more prevalent than ever. Aggregators have made it easier for customers to compare numerous car insurance companies by using one internet site, instead of contacting different companies directly. Marketing is important as it develops a company's brand equity, the value of the brand, which has been shown to increase their marketing communication effectiveness, thus promoting further growth opportunities (Huang and Sarigöllü, 2012). Within the internet environment (online) the company's website presents the first contact between the customer and the insurance company, the 'customer experience' (Dayal *et al*., 2000), but aggregators have the ability to remove this experience as they become the first point of contact.

Changes to distribution channels have been researched but tend to focus on companies utilising the internet for customers to contact them directly (Huang and Swaminathan, 2009; Pfiel *et al*., 2007; Wolk and Skiera, 2009) as well as the impact of insurance companies using a direct channel on a brick and mortar intermediary (Bouwman *et al*., 2005; Hoyt *et al*., 2006; Pfeil *et al*., (2008). This research explores how companies should expect aggregators to affect the channel choice and pricing strategies. Price comparison sites display the price a customer would get if they contacted the insurance company directly, so the price strategy of contacting the company directly would need considering. Within the UK, car insurance customers on the internet mainly have a choice of two avenues to get a quote, either via the company directly or through a price comparison site (as long as the company is displayed on a price comparison site). The main selling point of price comparison sites is that they can produce many quotes as quickly as a customer going to one company directly. If a customer chooses the price comparison site route and the insurance company is absent, then the company will lose out, but if they are present on an aggregator this could lead to cannibalisation.

There is extensive research regarding marketing a company online and the effect of price comparison sites, but these price comparison sites tend to be restricted to household goods (Papatla and Liu, 2009: Waldfogale and Chen, 2006). Financial price comparison sites are more complex, as they require the customer to input a lot of information about themselves, sometimes personal, so that an appropriate insurance quote is generated. The literature for price comparison sites within the UK insurance industry is sparse. The research located by this researcher notes Robertshaw (2011), who investigated customer profitability for general insurance. Also, McDonald and Wren (2009) implemented a manual search of contacting insurance companies directly on the internet, instead of using a price comparison site, to find that price changes depending on the person's profile. The absence of such research within the UK car insurance market is the basis of this present study.

This study principally aims to create a marketing model for the adoption of aggregators within the UK car insurance. This model is achieved by reviewing previous literature regarding price comparison sites, across different industries and marketing strategies, and performing key customer modelling analyses based on insights from the literature.

The outcomes of the research will produce a deeper understanding for the UK car insurance company that will benefit future research, senior strategic marketing managers and possibly other industries considering adopting aggregators as part of their distribution channel.

The remainder of this study is organised as follows. Section 7.2 reviews the general literature on the key marketing tools available for customer retention. Section 7.3 discusses the statistical models used in this study. Section 7.4 describes the data used and the methodology for this study. Section 7.5 presents the empirical analysis with the marketing framework. Section 7.6 summarises the main results and concludes the study.

7.2 Review of key concepts

7.2.1 Distribution Channels and price comparison sites

The arrival of financial price comparison sites (aggregators) has changed the way UK insurance companies conduct their business. The UK car insurance market has had to adapt and change its marketing, so that companies appear more relevant for their customers contacting them directly and indirectly via aggregators. Aggregators reduce search costs for the customer and allows comparison of car insurance prices easier. If a company chooses not to join an aggregator, this decision would make it harder for the customer to compare prices, due to the increased search costs, so the company can still discriminate their prices (Ellison and Ellison, 2009). Even if a company decides not to join a price comparison site, this would still affect their marketing, as they would need to stop customers comparing their prices, as this could potentially lead to a loss in sales, if a cheaper price was found with another brand.

When a company concentrates its marketing into one contact channel, this means the focus is on the product. This focus can be strengthened, as the company is not altering its message to appease different channels simultaneously (e.g. in the UK Direct Line insurance will keep on using a direct method only route). This strategy can also make companies 'concentrate on the cheapest channel system for that product' (Coelho and Easingwood, 2008, p.38). This may be beneficial to the company, but unless there is an available budget to get their voice heard over all the other distribution channels this may lead to the company becoming smaller. Paradoxically, then, a company with a strong marketing message may suffer financially if it is unwilling to extend its message to encompass different channels.

The arrival of aggregators does not necessarily mean that the car insurance company should neglect its traditional marketing strategies. Customers continue to contact the car insurance company directly, so a strategy needs to be put in place that 'co-ordinates marketing activity across individual customers, channels and marketing programs' (David Sheppard associates, 1999, p.71) The car insurance company needs to be aware that customers are used to using multi-media platforms, which will enable the customer to contact the company in numerous ways, e.g. phone, mobile phone, internet, internet on mobile phone, aggregator. The marketing department has to be aware that a customer viewing an advert on a bus can quite easily contact the company, as easily as a customer at home reading a magazine or watching television, so all avenues of contact need to be considered.

Price comparisons sites like to retain the companies on an exclusive basis; that is, that type of car insurance company is only available at this website. It has been shown 'that intermediaries do not like their suppliers to engage in multi-channel distribution' (Coelho and Easingwood, 2008, p.38), especially if the companies adopt a different pricing structure for each individual intermediary. The more exclusive companies an intermediary possesses, the better placing it has in the market. On the other hand the main consequence of using a single channel for the insurance company is that they are limiting their distribution. By using various distribution channels, including aggregators, a company can use the advertising of the aggregator to promote its own product and make contact with potential customers that they normally would not reach

7.2.2 Car insurance strategies within the UK

Different car insurance companies will apply different benefits, or remove certain benefits as customers 'balance the benefits of the purchase against the costs' (Grewal *et al*., 1998, p.56). In this instance the choice is with the customer to consider what benefits they would appreciate the most, e.g. car breakdown cover, accidental cover and for what cost. The effect of combining numerous products together is sometimes referred to as bundling. Bundling numerous products together makes it more difficult for customers to compare products between companies and so become more profitable (Linde, 2009). Hence, it is the insurance company that has to devise the recommended strategy. The company needs to consider what bundles should be most attractive to a particular type of customer and with advances in the management of aggregator channels, a company should be able to use price comparison sites to tailor different packages for the customer. In both scenarios, both the company and the customer need to consider the relevance and appropriateness of the product.

Aggregators can rank the product by price, or by benefits and this has led to greater comparison. Companies based at the 'top end' of the search tend to generate more sales (Arbatskaya, 2007; Haan and Moraga-González, 2011). For a company to appear at the top of an aggregator based on price, they can 'create an inferior quality good that can be offered at a very low price' (Ellison and Ellison, 2009). This could leave the customer with a very basic product that does not meet their needs. Within a product ranking environment the insurance company needs to 'refine the needs of users and match them with the appropriate product' (Laffey, 2010, p.1951). This produces a conundrum for the insurance company as if customers wish to purchase car insurance with all the benefits, then this will inevitably lead to a high price product which could be harder to locate if all the low priced, basic cover policies are shown first. Aggreators increase price transparency, but lower distribution costs, so they can sometimes be seen as both a blessing and a hindrance.

Without any additional benefits to a car insurance price, the price shown would just be the price attached to the potential customer. As well as customer risk being an attribute to price, McDonald and Wren (2009) also found that a customer's ability to search for different insurance companies affects their price. This demonstrates that companies use acquisition channels for their pricing behaviour (Brynjolfsson and Smith, 2000). Car insurance companies tend to advertise that if customers go to their company via their website, they would make a saving and with comparison sites on the internet, this behaviour should be expected. One of the main stipulations for an insurance company to be on an aggregator site is that the price should reflect the price of contacting the insurance company directly, which could lead to an increase in prices to accommodate aggregator payments.

When customers visit a price comparison site they have a perceived cost of the product. Prices within this bracket are deemed plausible, but prices outside this bracket, including cheaper prices, could act as a deterrent (Alford and Engellhand 2000). Customers have a budget and like to weigh up benefits with the price. Different customers want different benefits e.g. some customers may prefer to spend more money on their car insurance if they are guaranteed that all their calls are UK based. This demonstrates the importance of marketing and the message companies need to convey.

When a customer needs to contact their insurance company, they need to know that they are being dealt with by trained, professional and experienced staff so employee selection/training and motivation is important. Delivering excellent customer service 'relates to and determines the level of satisfaction the customer receives' (Murphy *et al*., 2006, p. 78). When a customer contacts their car insurance company, it could be to report an accident and make a claim. The claim process can be quite a traumatic time for a driver and having motivated trained staff delivering excellent service may ensure the retention of the customer, even though their premiums may rise. In this instance, a good customer service could also improve personal referrals.

7.2.3 Social network sites

Aggregators belong on the internet, where customer reviews of a particular brand are prevalent. Word of mouth (WOM) marketing has been shown to have a stronger effect on acquisition, than just using traditional marketing techniques (Trusov, 2009: Degraffenreid, 2006 and Libai *et al*., 2012). Traditional marketing, (print, outdoor, TV) is still relevant but within an aggregator environment, insurance companies need to be aware of the influence of social media sites. Traditional marketing can be used to give a brand its brand equity which 'represents the degree to which a brands name alone contributes value' (Leuthesser *et al*., 1995, p.57). Using a company's brand equity, alongside WOM, may generate more positive reviews and can help limit any collateral damage encompassed from negative reviews.

Social networking sites are a powerful way of spreading relevant information to customers and potential customers. It is easy for these to remove or 'cull' unwanted friends or companies from their page, so companies need to have 'something to say that is meaningful, useful, interesting and has personal value' (Tapscott, 2009, p.81). A company should try to get their customers to share positive company messages with their friends. When a customer recommends a product to a friend, these friends tend to convert more efficiently (Degraffenreid, 2006: Trusov *et al*., 2009). This scenario indicates that the car insurance company must keep its updates relevant and interesting, if they wish to convert extra customers that they may not potentially have reached before.

The effect of social network sites cannot be directly analysed, but personal recommendations can be which gives rise to the following hypothesis:

H1: There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate

This hypothesis will tell the company how much external factors affect their aggregators' acquisition rates. As previously noted, the UK car insurance industry is different to other industries as it is a legal requirement. Each personal recommendation is a strong marketing tool, which also gives the company a slight insight on how they are conducting business. If the company is working positively for the customer, then this also generates positive feedback and vice versa.

7.2.4 Advertising in an aggregator world

Aggregators can potentially give the insurance company extra customers, so if the company joins an aggregator the marketing strategy needs to be amended. Within the marketing department, 'selling via intermediaries requires that marketing effort is directed at both the intermediaries and the end customer' (Harrison, 2000, p.91). The insurance company would prefer the customer to contact them directly instead of via an aggregator, as this would save them costs, but with aggregator marketing budget greater than the insurance company budget, the marketing strategy needs to be compromised. The insurance company may have to demonstrate with its marketing that it is a premium product at a reasonable price and that 'customers pay lower prices in aggregate, but not all customers are better off' (Thomas, 2012, p.38). This scenario would hopefully make more customers contact the company directly, but if the customer does purchase via an aggregator, they should expect a high-quality product.

Certain customers have particular shopping habits, where they continue to purchase their goods in the same manner constantly. Marketing can be used to change a customers' shopping habits or their entrenched buying behaviour, which when 'people get used to buying certain products through particular intermediaries and have an inbuilt inertia to change' (Wilson *et al*., 2008, p.531). The other side to this argument is to stop customers going to aggregators to change their insurance company. Aggregators need customers to change their insurance each year to generate the most profit. Marketing techniques have the potential to stop customers leaving their current insurance company and from using aggregators to review their renewal prices. This scenario gives rise to the next hypothesis, again with the previous scenario recalled for easy follow through:

H1: There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate

H2: The more money spent on advertising, the more customers will contact the company directly instead of going to an aggregator

As mentioned predominantly in chapters 2 and 4, marketing provides the customer information of what service they should expect from the company and what they sell. For every sale made via an aggregator, the car insurance has to pay them. If the insurance company can stop customer using aggregators with their own marketing budget, this should save the insurance company money. Aggregator advertising is not considered for this hypothesis, as it is reflects the factors that the UK car insurance company can influence.

7.2.5 Internet advertising

Internet and search engine advertising is very important in this aggregator environment. A customer who searches for the company on a search engine using the company's name (natural search) demonstrates that the marketing is effective. To become prominent using a generic term on a search engine, e.g. car insurance, the company has to become prominent/top, where the competition is fierce (Haan, and Moraga-González, 2011). Customers use search engines, as they can provide customers with direct access to the products and information they require and so that firms can find their target (Shih *et al*., 2012). This could potentially lead to brands spending a lot of money to appear first on a search engine and neglecting other marketing techniques that would lead customers to search for the actual brand instead and reduce costs. This implies that focussing advertising on one particular medium can be counterproductive and expensive.

As well as internet search, the internet contains banner ads, which directs the customer to the website when the customer clicks on them. The issue with banner ads is their lack of branding capability (Vries, 2012). Banner ads act as a portal between websites, but they tend to lack content. Only if other marketing strategies have been used, to give the brand recognition, can banner ads be truly effective. This does not mean that banner ads cannot change their message. Any messages on banner ads need to be truthful as over two thirds of customers would not buy a product if they found the marketing message to be untrue (Tapscott, 2009). This demonstrates the need to have constant clear, concise and truthful messages in all marketing communications. Any lapse in professionalism could spread through other acquisition channels.

This section prompts another hypothesis to test whether own marketing practices have moved customers away from aggregators online as customers search for the insurance company directly. This gives rise to the next hypothesis, again with the previous scenarios recalled for easy follow through:

H1: There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate

H2: The more money spent on advertising, the more customers will contact the company directly instead of going to an aggregator

H3: An increase of internet marketing spend increases the likelihood of customers choosing direct web channel instead of using an aggregator.

As mentioned previously in this thesis, aggregators are web based tool and do not function off-line. This hypothesis tests that if the insurance company tries to compete with the aggregators in their own environment, on-line, this will this drive customers away from the aggregators.

7.2.6 Direct Marketing

For direct marketing to be effective, information about the customer needs to be correct and relevant. Direct marketing can be used not just for contacting and transacting (Kotler, 1997) but to communicate 'high quality and high value in building customer loyalty (David Sheppard Associates, 1999, p.79). Direct marketing can be used to provide insight into the insurance company's strategy and how it wishes to perform for the customer. Done correctly, direct marketing can give the impression of the company being professional and reliable. The power of direct marketing can be extended to build relationships with the customer. The main issue with direct marketing in the UK car insurance industry is that it is an annual product and sending out direct mailing when the customer has just purchased their insurance from another company, could lead to wastage.

Direct marketing strategies prompt the following hypothesis, again with the previous scenarios recalled for easy follow through:

H1: There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate
H2: The more money spent on advertising, the more customers will contact the company directly instead of going to an aggregator
H3: An increase of internet marketing spend increases the likelihood of customers choosing direct web channel instead of using an aggregator
H4: Direct marketing increases direct sales but not aggregator %

Direct marketing can come in numerous forms, email, post, SMS or telephone calls. Hypothesis 4 tests whether contacting the customer directly, by post, will make customers contact the insurance company instead of using price comparison sites.

7.2.7 Relationship marketing

The marketing department needs to be aware of its customers and of the Paretos principle, also known as the 80/20 rule, where 80% of profits are generated from the top 20% of customers. Relying on new customers for company growth is short-sighted as long term customers can generate more profits for the company. Customer retention must not only be part of marketing but also ‘be integral to a company’s basic strategy’ (Reichheld, 1993, pp.64). The insurance market with in an aggregator environment has to try and stop profitable customers defecting and searching for cheaper quotes. Focusing on renewing profitable customers, through direct mailing offers and great customer service can help build the company with profitable customers.

The insurance market is split into numerous products and each product must be treated differently. Within insurance, the product and message must be integral as it has been shown that home insurance customers behave different to car insurance customers (Thuring *et al.* 2012). A home insurance customer may prefer tips and offers regarding home insurance and a car insurance customer may prefer car tips. This does not necessarily mean that the car insurance customer should just receive offers regarding cars, all customers are different and with car insurance being a legal requirement, offers for hotel trips or days out could be more beneficial for certain customers.

The effect on renewals within the marketing mix prompts the final hypothesis, is, yet again with the previous scenarios recollected for easy follow through:

H1: There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate
H2: The more money spent on advertising, the more customers will contact the company directly instead of going to an aggregator
H3: An increase of internet marketing spend increases the likelihood of customers choosing direct web channel instead of using an aggregator
H4: Direct marketing increases direct sales but not aggregator %
H 5: Non-DM and non-web (other) marketing affects the renewal rate

Finally, Hypothesis 5, considers the renewal rate. As shown in chapters 4 and 6, aggregators do affect the renewal rates. This hypothesis combines all of the other traditional marketing spend to see the effect it has on renewal rates. The different hypotheses proposed will provide useful insightful that will enable the development of a marketing framework.

7.3. Statistical models used in this chapter

For hypothesis 1, the effect of WOM on aggregator conversion rate, the aggregator conversion rate (ACR) needed to be calculated

$$ACR = \frac{AGG\ (sales)}{AGG(quotes)+AGG(sales)} \tag{7.1}$$

To test their influence, correlation statistics, using Pearson Chi-Square (PCS) test, were used. The PCS test calculates if two variables are totally independent or not. PCS is calculated using the covariance of the two variables, divided by the product of their standard deviations:

$$\rho_{x,y} = \frac{cov(X,Y)}{\sigma_x \sigma_y} \tag{7.2}$$

Where cov is the covariance, σ_x is the standard deviation of X and σ_y is the standard deviation of Y

If the WOM sales increases/decreases as ACR increases/decreases, then they can be considered as correlated. The closer the value is to one, the less independent WOM and ACR are, therefore we can measure how much influence WOM has on ACR. For Hypothesis 2, the effect of marketing on direct sales and aggregators, the marketing spend is the non-internet and non-DM spend. The marketing spend used in this scenario is more focused with the outdoor marketing, the branding of the company (TV, radio, outdoor, print). This will test whether an increase in the ambient marketing will make more customers contact the company directly, thus decreasing the ACR (Equation7.1). A graph was produced of spend against sales and ACR, by month. The ambient marketing spend and direct sales (non-aggregator) were indexed so that 1 is the average.

$$marketing\ spend\ (indexed) = \frac{marketing\ spend\ for\ month}{average\ all\ marketing\ spend} \tag{7.3}$$

$$direct\ sales\ (indexed) = \frac{direct\ sales\ for\ month}{average\ all\ direct\ sales} \tag{7.4}$$

Next, hypothesis 3, the effect of internet advertising against aggregators is to be investigated. Due to timings of gathering accurate data, the dates up to and including

March 2009 were used. A graph was produced detailing web spend, web sales and ACR (Equation 7.1), by month. The marketing spend and the number of sales was indexed, where 1 was the average, to track the performance more informatively.

$$web\ spend\ (indexed) = \frac{web\ spend\ for\ month}{average\ all\ web\ spend} \tag{7.5}$$

$$web\ sales\ (indexed) = \frac{web\ sales\ for\ month}{web\ all\ direct\ sales} \tag{7.6}$$

Concerning hypothesis 4, the effect of direct marketing on aggregators entails plotting a graph of DM spend against ACR, by month. The DM spend was indexed, where 1 equals the average, to track the performance more informatively.

$$DM\ spend\ (indexed) = \frac{DM\ spend\ for\ month}{average\ all\ DM\ spend} \tag{7.7}$$

Finally, hypothesis 5 investigates the effect of non-DM and non-web (other) marketing spend on the customer renewal rates.

$$Renewal\ Rate = \frac{Number\ renewed}{Number\ renewed + Number\ not\ renewed} \tag{7.8}$$

A graph was produced of the marketing spend against the renewal rate, by month. The ambient marketing spend was indexed, where 1 is the average, to track the performance more informatively (equation 7.7)

7.4. Empirical analysis and results

7.4.1 Data description

To test the different hypotheses, data from an established UK car insurance company was used. The insurance company had joined a price comparison site and has been amending its marketing budget, due to the costs of aggregator sales using the marketing budget. The data set provided from the insurance for this analysis is at monthly levels for the period covering May 2007 to August 2009 inclusively, containing:

- Month of insurance quote enquiry
- Aggregator conversion rate (ACR)
- Word of mouth (WOM) sales (indexed)
- Ambient marketing spend (indexed)
- Direct (non-aggregator) sales (indexed)
- Web sales (indexed)
- Web spend (indexed)
- Direct mail quotes (indexed)
- Direct mail marketing spend (indexed)
- Percentage of customers retained

Table 7.1: Descriptive statistics of data used

Value	mean	Std. dev.	max	min	skewness	kurtosis
ACR	1.51%	0.58%	2.47%	0.72%	0.09	-1.50
WOM	2603	1543	5273	642	-0.02	-1.60
Ambient	1.03	1.01	4.18	0.10	1.52	1.91
Direct	1.00	0.79	3.22	0.20	1.10	0.93
web	1.00	0.66	2.60	0.27	0.71	-0.20
webspend	1.00	0.64	2.13	0.08	-0.02	-1.31
DMquote	1.00	0.71	2.71	0.24	1.05	0.17
DMspend	1.00	0.75	2.86	0.18	0.67	-0.41
Retained	69.93%	1.60%	73.00%	67.00%	-0.04	-0.80

These date periods are relevant as they cover the dates when the company joined a price comparison site, and after. Between May 2007 and August 2007 the company was in its testing stage to make sure that its systems and infrastructure could manage this new distribution channel. From September 2007, the company was fully integrated with the aggregator.

The media source specification for the sales and spend is gathered from either:

- The customer answering a question when gathering a quote
- Being completed automatically if a specific web link is clicked
- Being automatically completed if quote arises via aggregator.

The media source specifications have been broken down into the following segments:

Table 7.2: Media source specifications

Source	Example
Web	Banner ads and search engine
Direct Mail	Postal and email
Word of Mouth	Personal Referrals
Aggregator	Price comparison site
All Other (ambient)	TV, Radio, Print, Outdoor

7.4.2 Framework development

To understand the new environment of aggregators marketing departments need to develop a new strategic framework. Insurance companies need to understand the current situation as their marketing decisions are based on their models (Leeflang and Wittink, 2000: Buckin and Gupta, 1999). Reviewing the results from the hypothesis, a detailed informative framework can be constructed. This framework is shown in Figure 7.5 below.

7.5. Results

7.5.1 Hypothesis results

Hypothesis 1

Table 7.3: Pearson Correlation statistics of ACR and WOM

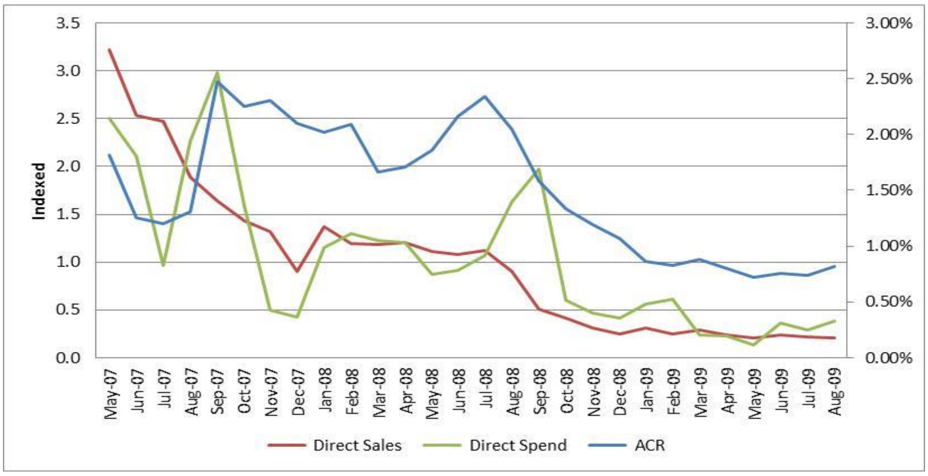
	ACR	**WOM Sales**
ACR	1	0.76771 (<0.001)
WOM Sales	0.76771 (<0.001)	1

From table 7.3, it can be noted that WOM and ACR are highly significantly correlated with each other, which means that WOM influences aggregator conversion rate. These results agree with Degraffenreid (2006) who also found that people who refer products influenced customers' purchasing habits and with the market response modelling results in Chapter 4.

Hypothesis 2

From Figure 7.1, it can be shown all the variables are significantly correlated, and that marketing spend has dropped since the arrival of aggregators, which is due to aggregator spend using the marketing budget. The graph shows that a reduction in marketing spend affects both the direct sales and ACR negatively, also there is a lag effect, so the impact is not observed initially. The results correspond with Leuthesser *et al* (1995) who found indirect measures for brand equity to have an influence on choice. With less brand awareness due to a lack of marketing, this conversely reduces the ACR and sales.

Figure 7.1: Comparing direct sales and ACR by marketing spend



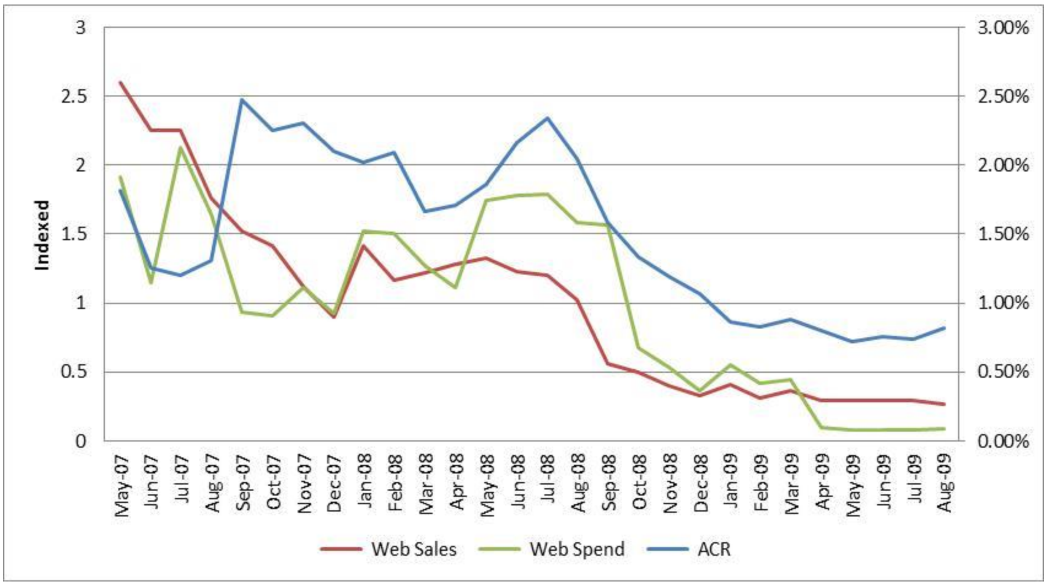
Key: indexed refers to the indexed figures as detailed in section 7.3, right axis is the ACR

Table 7.4: Correlation statistics of direct sales, ACR and marketing spend

	Direct sales	ACR	Marketing spend
Direct Sales	1	0.487 (0.009)	0.710 (<0.001)
ACR	0.487 (0.009)	1	0.470 (0.011)
Marketing spend	0.710 (<0.001)	0.470 (0.011)	1

Hypothesis 3

Figure 7.2: Web sales and ACR by web marketing spend



Key: indexed refers to the indexed figures as detailed in section 7.3, right axis is the ACR

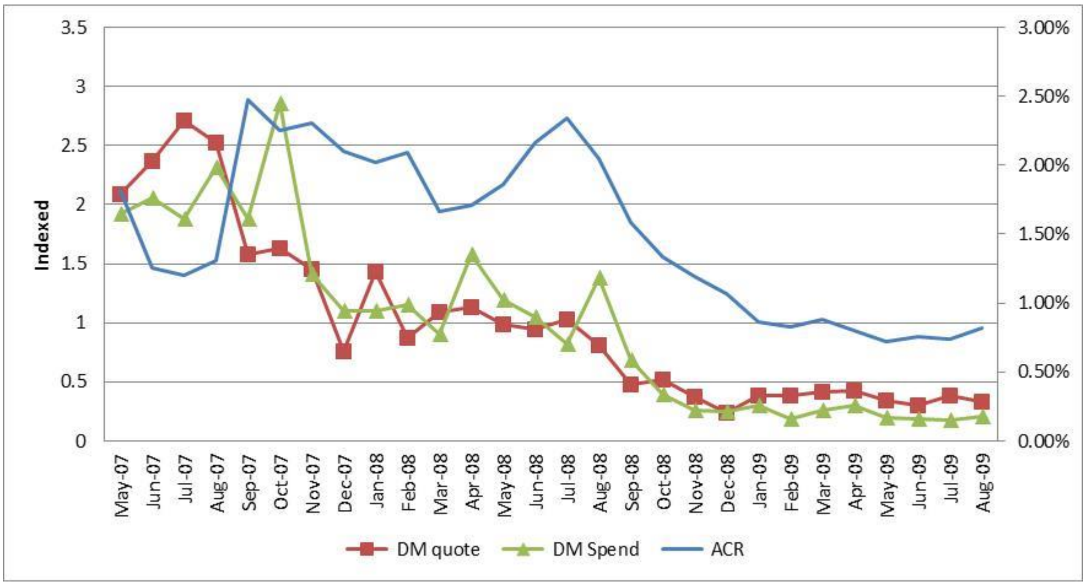
Table 7.5: Correlation statistics of direct sales, ACR and marketing spend

	Web sales	ACR	Web Marketing spend
Web Sales	1	0.538 (0.003)	0.793 (<0.001)
ACR	0.538 (0.003)	1	0.679 (<0.001)
Web Marketing spend	0.793 (<0.001)	0.679 (<0.001)	1

From Figure 7.2, it can be perceived that as web spend decreases, so do web sales and the ACR and this relationship is statistically significant. In this scenario (H3), web spend does not take customers away from the aggregator. This corresponds with Ranfaswamy (2005) who found that only by observing customers' behaviour across channels can we improve understanding of the customer. Aggregators are a web-based tool, but it seems customers who come to the company directly via the web, may not use aggregators.

Hypothesis 4

Figure 7.3: The effect of DM spend on DM quotes and ACR



Key: indexed refers to the indexed figures as detailed in section 7.3, right axis is the ACR

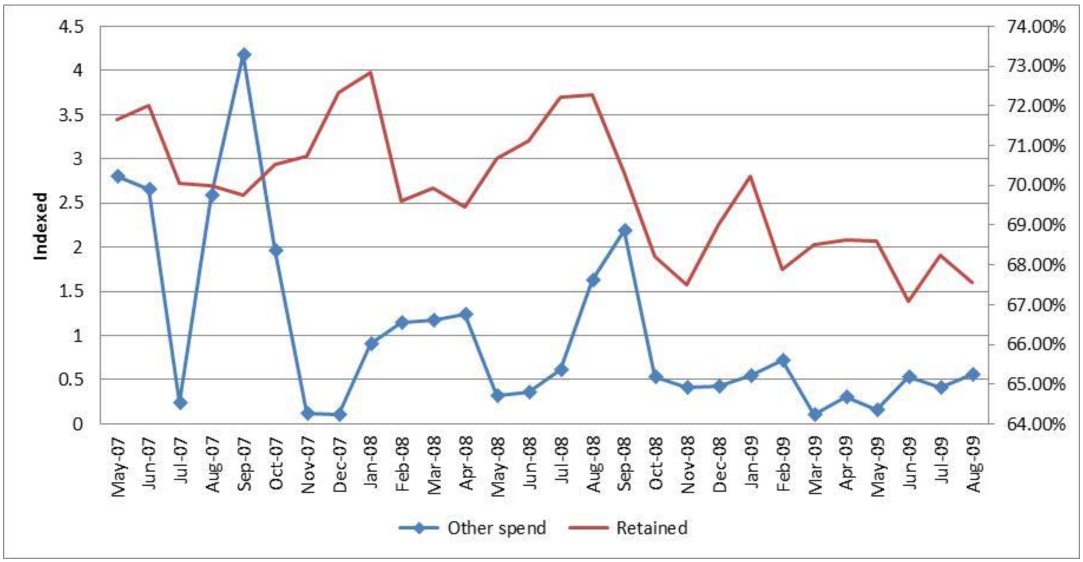
Table 7.6: Correlation statistics of DM quotes, ACR and DM marketing spend

	DM quote	ACR	DM spend
DM quote	1	0.516 (0.003)	0.894 (<0.001)
ACR	0.516 (0.003)	1	0.678 (<0.001)
DM	0.894 (<0.001)	0.678 (<0.001)	1

From Figure 7.3, it can be noticed that DM does not directly affect but only moderately influences ACR, but does influence DM quotes. Direct marketing, in this instance, behaves differently to outdoor marketing, due to the small effect it has on ACR. Figure 7.3 also demonstrates that DM marketing does affect the DM channel, which slightly differs from Coelho and Eastwood (2003), who found that we cannot predict channel preference with confidence, albeit only in a limited campaign metric.

Hypothesis 5

Figure 7.4: Marketing spend against renewal rates



Key: indexed refers to the indexed figures as detailed in section 7.3, right axis is the retained rate

Table 7.7: Correlation statistics of DM quotes, ACR and DM marketing spend

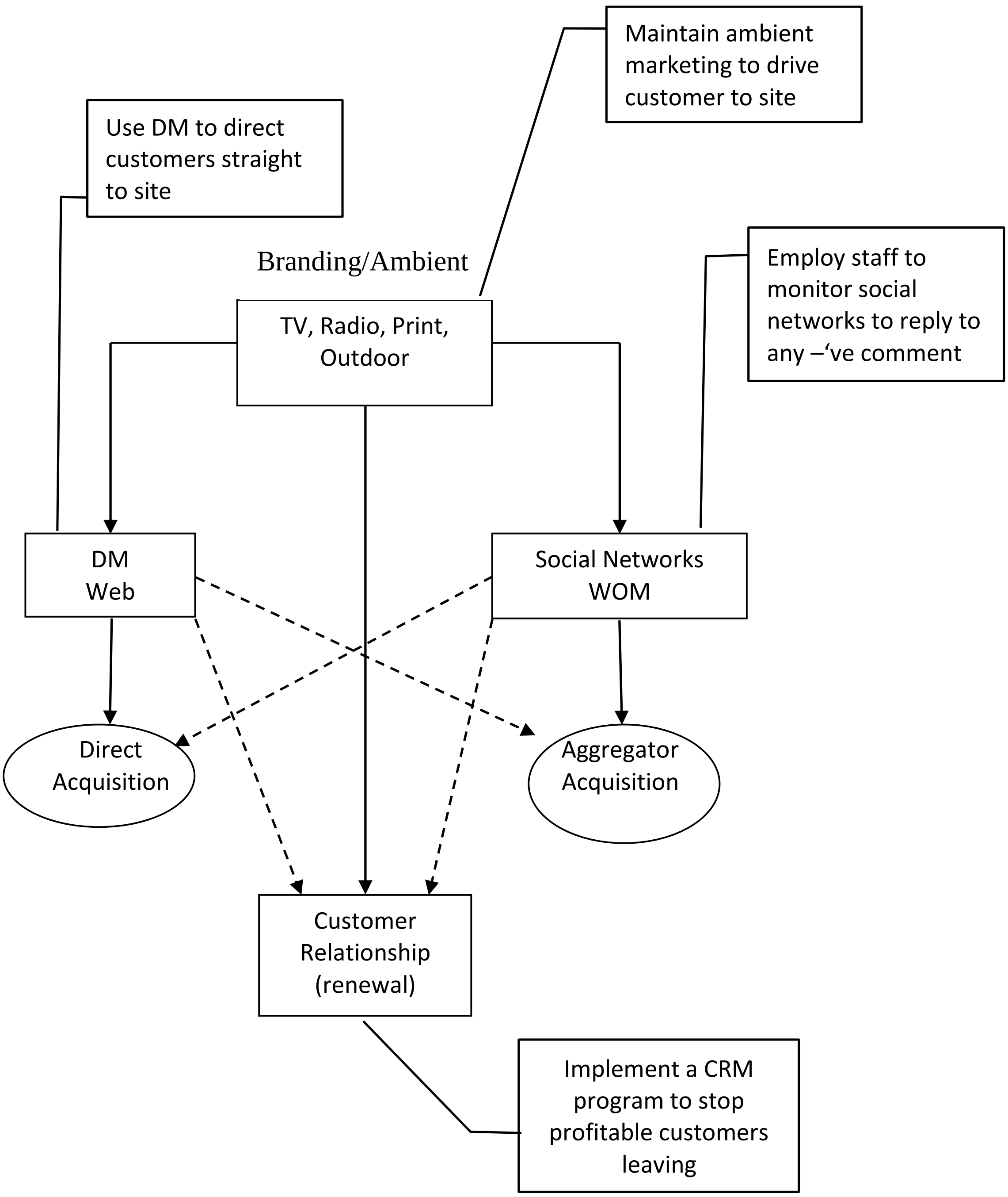
	ACR	WOM Sales
ACR	1	0.453 (0.016)
WOM Sales	0.453 (0.016)	1

Figure 7.4 demonstrates some instability, but the drop in marketing spend along with renewal rates can be observed. The company had joined an aggregator in May 2007 and until September 2008, renewal rates had remained quite constant, but the drop in marketing spend has coincided with a drop in renewal rates. The results correspond with Leuthesser *et al* (1995), who found that less brand awareness, due to a lack of marketing, reduced retention rates.

7.5.2 Development of a marketing framework

From the different hypotheses constructed it can be shown that DM (H4) and website (H3) by themselves moderately influence aggregators conversion rates. What has been discovered is that WOM (H1) and a combined marketing budget (H2) will affect aggregator conversion rates. This demonstrates that web marketing and DM should be primarily used for customers to contact the company directly, but its secondary influence should not be disregarded, as the person who received the direct mail, may pass this to one of their friends and thus become a word-of-mouth customer. Also renewal rates (H5) can be affected by marketing, so this needs consideration when producing the framework.

Figure 7.5: Marketing framework



Key —— = Primary ---- = Secondary

Figure 7.5 demonstrates the primary and secondary actions regarding marketing in an aggregator environment for UK car insurance companies. The branding marketing affects all acquisition channels as well as renewals. Branding marketing along with DM and web marketing contribute primarily for the direct channels, but DM and web can also act as secondary effect by increasing the brands knowledge Ng (2004).

For aggregator acquisition, social networks/WOM do have an influence, which agrees with Trusov *et al*. (2009), who also found that social networks/WOM have a strong effect on acquisition. This also implies that the secondary effect on direct acquisition should not be dismissed. Finally, from graph 7.4, it is seen that traditional marketing does have an effect on renewal. The framework also points out that the company should not use social networks and DM just for acquisition; they should also be used to help retain customers. Sharp *et al*. (2002) consider insurance to be a "subscription-type", which has demonstrated strong brand loyalty that is able to satisfy the customers' needs. The annual cover purchased for car insurance can give such companies a chance to build a relationship with their customers throughout the year.

7.6. Summary and conclusion

7.6.1 Summary

The different hypotheses produce subtly different insights into marketing and customer behaviour during the re-intermediation for the UK car insurance industry.

Table 7.8: Hypothesis results

Hypothesis	**Result**
H1: *There is a positive relationship between word-of-mouth (WOM) sales and the aggregator conversion rate*	True
H2: The more money spent on advertising, the more customers will contact the company directly instead of going to an aggregator	Moderately true
H3: An increase of internet marketing spend makes customers choose direct web channel over of aggregator	Moderately true
H4: Direct marketing increases direct sales but not aggregator %.	Moderately true
H5: Non-DM and non-web (other) marketing affects the renewal rate	True

A combination of the different marketing techniques can influence the acquisition rates across all the different channels and can help retain customers. Also, the marketing framework must acknowledge the secondary effects, as well as primary effects, of the different media campaigns.

7.6.2 Conclusion

The main purpose of this study was to increase the understanding of the marketing mix within an aggregator environment for the UK car insurance industry and by building a framework to represent the challenges of the marketing strategies (objective 3). To develop this framework, impact analysis was conducted by proposing certain hypotheses which covered direct channel acquisition, aggregator channel acquisition and customer retention.

Firstly, this research found that word of mouth advertising has a strong effect on ACR, whereas DM and web advertising mainly effect direct channel acquisition. Also this research found that using ambient marketing (outdoor, TV, print radio etc.) impacts ACR retention and direct sales.

This research proposes that aggregators should be treated as another acquisition channel and that a multi-channel strategy should be used to its full advantage instead of being viewed as a hindrance. When the internet commerce first occurred, original forecasts suggested a complete change to how a business operates, but Harris and Coles (2004) found that a combination of old and new practices worked more efficiently.

This research also recommends that web and DM advertising ought to become more involved in marketing strategy and be considered for use with customer retention. The use of multi-channel communication should be used to build on the customer/company relationship (Payne and Frow, 2005). For a company considering costs, as postal direct mail often produces a wastage (Ng, 2005) due to its low response rate of roughly 2 % (Stern and Priore, 2000), emails can be used to provide more direct access to the customer.

This work contains some important implications for marketing strategy teams. The different hypotheses tested gave insight into what marketing practices impact different scenarios. Also, the framework offers managers a tool to understand how their marketing budget is being used. Finally this research proposes a new direction for DM and web marketing to help with customer retention for the company.

7.7 Limitations and Further Research directions

The limitations of this research should also be considered. The data only comes from one insurance company, so if another company is using a different marketing budget strategy, this presents a limitation. Also this research only considers companies using aggregators, and does not consider companies that have not joined an aggregator.

Another limitation to consider is the personal recommendation statistic. There is not a clear way to note if the recommendation came from friends outside or inside social networking sites. In general, further research would be required to establish whether these results can be applied to other companies.

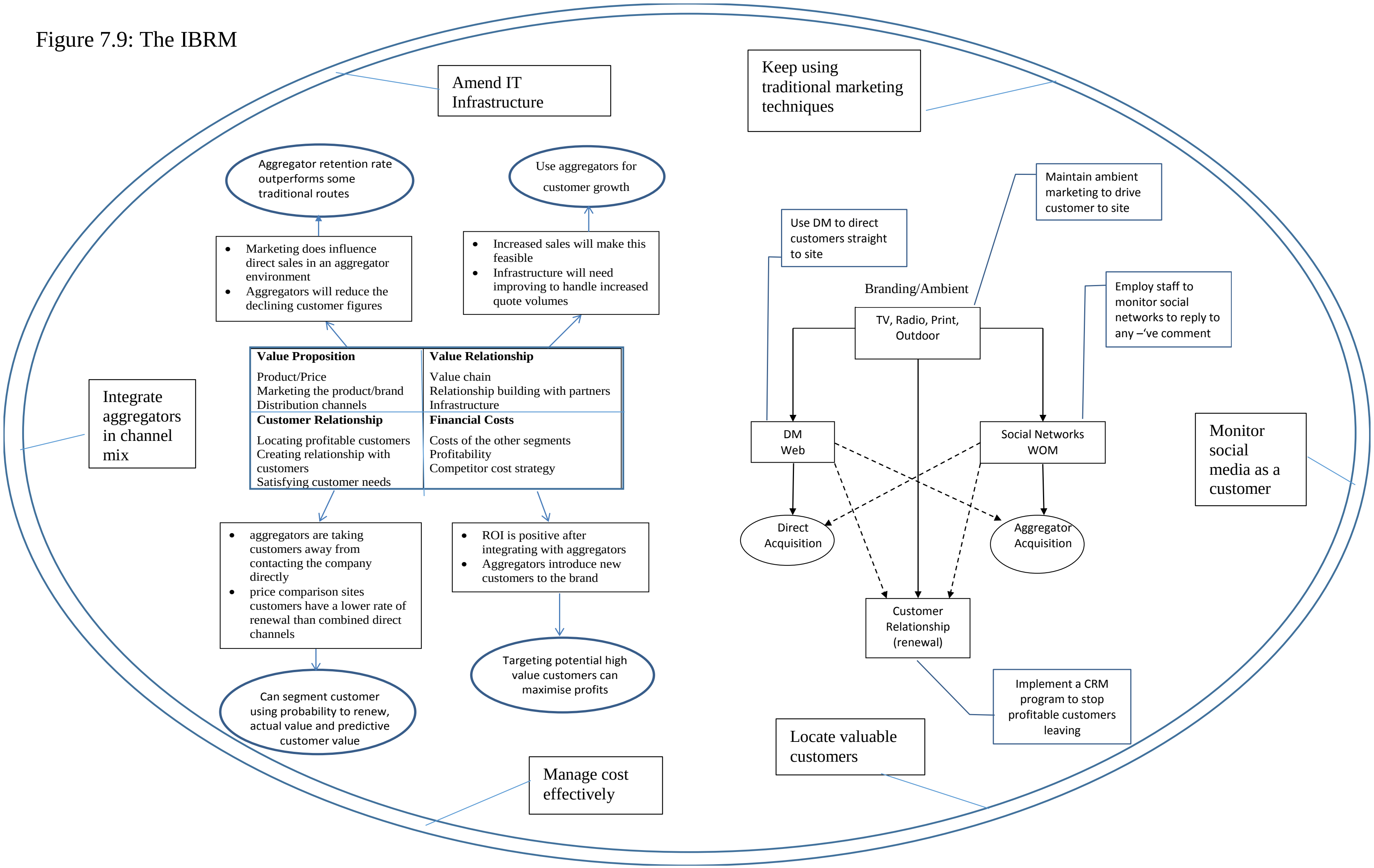
When reviewing the ambient marketing techniques, the research does not take into account any regional variations that may occur with different types of marketing being applied.

The main goal of this research was to provide a deeper understanding the impact price comparison sites have had on traditional marketing within the car insurance environment. This was completed by analysing different hypotheses and developing a marketing framework. Thus, this is among the first to review traditional marketing practices on a price comparison site within the UK car insurance environment. In terms of the systematic development of proposed IBRM framework for car insurance marketing, Figure 7.9 below captures all the model elements from Chapter 5 to Chapter 7 and constitutes a major contribution of the research to knowledge, as mentioned earlier.

Figure 7.9 (the IBRM) contains a combination of figures 6.14 and 7.5, surrounded by the key messages from this research to keep the case company profitable:

- Integrate aggregators in channel mix
- Amend IT structure
- Keep using traditional marketing techniques
- Monitor social media
- Locate valuable customer
- Manage costs effectively

Figure 7.9: The IBRM



Chapter 8: Conclusions and recommendations

8.1 Introduction

This study's main aim was to understand the effects of re-intermediations within the UK car insurance industry and how the companies can adapt to this new environment. Price comparison sites (aggregators) belong to the cybermediation scenario, which enables the customer to compare the prices of different car insurance companies. Their impact within the insurance environment cannot be considered minor. They provide valuable information to the customer as well as saving them the time of contacting numerous companies individually. The UK car insurance industry is fully aware of this new environment, thus this research uses strategic marketing, customer relationship and market response modelling to explore the effects of aggregators on insurance business success.

This research is the first study to fully investigate the effects of price comparison sites based on core business metrics, namely marketing, quotes, sales and customer retention within the UK car insurance industry. An understanding of these characteristics is important to the industry given the lack of current research in this area.

The research is also the first to develop an integrated business model and a marketing framework within this new environment (the IBRM). This will prove useful for the car insurance industry and other insurance providers (home, van, pet). It will also be useful for other industries that are about to become involved in an aggregator environment, and for senior members of staff within the car insurance industry.

8.2 Main results of the research

The main results of the study are summarised below.

8.2.1 Long and short term effects

A vector autoregressive (VAR) model was developed in chapter 4 using different channel acquisition rates and marketing spend. The key findings from this model are:

a) Customers acquired via other direct routes, have a long-lasting positive effect on word-of-mouth (WOM), aggregators, spend and on direct routes, but a negative effect on retention and win-back
b) WOM has a strong positive effect on future WOM acquisition rates, no effect on marketing spend, negative effect on win-back, retention and aggregators. WOM also had an exceptionally negative effect on the other direct routes
c) Win-back has erratic short term effects, but the long-term effects demonstrate a strong negative effect on the other direct routes, a negative effect on retention rates, a slight negative effect on aggregator rates but strong positive WOM effects.
d) Retention has a strong effect on other direct routes and retention, a slight positive effect win-back and aggregators, but a slight negative effect on spend and a stronger negative effect on word-of-mouth
e) After 12 months, aggregators have a positive effect on retention and other direct routes ratios and retention but a negative effect on WOM spend and a very negative effect on win-back.
f) Marketing spend has a strong effect on other direct routes, WOM, aggregators, but a negative effect on win-back and a strong negative effect on retention

The results presented above provide an insight into the mechanics of how re-intermediation relates to strategic marketing planning and its implementation via the marketing mix. The findings contributed in understanding the repositioning of the case company with regards to its future growth and profitability

8.2.2 Business model scenarios

Following chapter 4, the investigation of the holistic impact of aggregators, chapter 5 considered the business model. Numerous scenarios were answered in chapter 5, to address the impact of aggregators:

a) If a company does not join a price comparison site then it should still expect a good return on investment
b) Not joining a price comparison will cause a reduction of the number of customers contacting the company site
c) Not joining a price comparison site will lead to a reduction in customer retention

d) It is worth investing in extra resource and expenditure to enable aggregators in the distribution mix

Again, the different scenarios provided insight into the development of the business model. The business model showed that even though a car insurance company can still operate in a price comparison environment, by not including comparison sites in their distribution channel they should expect reduction in business.

8.2.3 CRM development

Chapters 4 and 5 revealed that including a price comparison site in the distribution list will gather more quotes and produce stronger retention rates, but this does not mean that the UK car insurance company should not try and reduce the number of customers using price comparison sites. Aggregators generate their money from the UK car insurance company for every sale made via the aggregator.

A significant discovery, in chapter 6, found that customers that have been with the company three years or more are, on average, more valuable than customers in their first year, even if the customers in the three-plus years segment have claimed. This insight demonstrated that customer retention can generate significant profits for the insurance company.

To locate these customers, chapter 6 produced a framework, based on actual customer value, predicted customer lifetime value and likelihood of them renewing. Numerous statistical models and data mining techniques were used:

- Neural networks (customer lifetime value & likelihood to renew)
- Decision trees (customer lifetime value & likelihood to renew)
- General linear models (customer lifetime value)
- Quantile regression (customer lifetime value
- Logistic regression (likelihood to renew)

with the following techniques being championed:

a) Quantile regression to predict customer lifetime value
b) Logistic regression to calculate the likelihood of customers renewing.

Actual customer value was added so that the company could consider whether it is worth renewing customers that may have cost them a significant amount of money previously.

8.2.4 Marketing framework

Chapter 7 uses the knowledge and findings from chapters 4, 5 and 6 to create a new marketing framework for the industry. Chapter 7 answered different hypotheses to maximise the marketing strategies:

a) The greater the word-of-mouth sales, the greater the aggregator conversion rate
b) Marketing spend will not stop the customers contacting the company directly instead of going to an aggregator
c) Increased internet marketing spend does not make customers choose direct web channel over aggregator
d) Direct marketing increases direct sales but not aggregator %
e) Ambient marketing affects the renewal rate

These findings enabled the production of a marketing framework that incorporated the different channels into their primary and secondary functions.

8.3 Business implications for the case company

Car insurance is a legal requirement in the UK, so all drivers in the UK must have a car insurance product. In 2011 the UK car insurance industry received £13.3 billion in premiums and insured 23.8 million private vehicles (ABI, 2012). These figures demonstrate the significant size of the industry and the business implications this research contains.

Traditionally, gathering car insurance quotes is a time consuming process, but aggregators have reduced this timely process. Aggregators have made it easier for car insurance quotes to be compared against each other, whether on price or additional benefits. In this section we link the key ideas and results from chapter 4 to 7.

Firstly, the market response model findings agree with Pauwels and Neslin (2008) that adding a new channel does affect customer retention. Aggregators make it easier for the consumer to shop around and get the best deals when considering car insurance, but the results provide insight that customers who used aggregators initially, may be more inclined to renew. As aggregators tend to have a bigger marketing budget than car insurance companies, the channel choice of the customers tends to be the aggregator site, which is consistent with Ansari *et al* (2008).

Customers who came to the insurance company via aggregators tend to have lower renewal rates than those who came through direct channels, but the trend of the direct channels renewal rate was decreasing pre and post joining aggregators. When switching costs are set higher it is easier to retain customers (Gronhaug and Gilly, 1991), but with reduced switching costs, the company should expect a lower retention rate. The lower retention rate with aggregators could be due to the fact that they are a web based tool, which is consistent with Ansari *et al*. (2008).

Another issue that would need to be considered with the renewal rates is the relationship building process with the customer. If the first contact with the insurance company is via an aggregator and the person then purchases via the aggregator, then this limits the possibilities of building that initial relationship with the customer. This is consistent with Coulter and Coulter (2002)'s research.

When considering different statistical modelling techniques and data mining tools, it was discovered that quantile regression performs the strongest when calculating CLV, which corresponds to the research conducted by Benoit and Poel (2009). Secondly, Winsorized data models perform better at the extremes when considering hit rate analysis, but for general linear models only. Finally, these techniques improved on Malthouse and Blattberg's (2005) findings, which found that out of the top 20% of customers, 55% would be misclassified as poor performers.

This research proposes that aggregators should be treated as another acquisition channel and that a multi-channel strategy should be used to its full advantage instead of being viewed as a hindrance. When the internet commerce first occurred, original forecasts suggested a complete change in how a business operates, but Harris and Coles (2004) found that a combination of old and new practices worked more efficiently.

The research also recommends that web and DM advertising ought to become more involved in marketing strategy and be considered for use with customer retention. The use of multi-channel communication should be used to build on the customer/company relationship (Payne and Frow, 2005). For a company considering costs, as postal direct mail often produces a wastage (Ng, 2005) due to its low response rate of roughly 2 % (Stern and Priore, 2000), emails can be used to provide more direct access to the customer.

Finally, Figure 7.9 presents an integrated car insurance marketing re-intermediation model (IBRM) which combines key insights from Chapters 4-5 of the thesis. The nature of contributions of this research to knowledge embodied in the model is further discussed below.

8.4 Summary of contribution of the research to knowledge

As previously mentioned in this thesis, there is a dearth of research when considering the UK car insurance market, especially within a price comparison environment. One main theoretical contribution to knowledge is the creation of the IBRM which links a car insurance business model with a CRM strategy. The IBRM incorporates a business model to assess the way in which a firm can combine insights from value propositions for different customer segments, customer life time values, effects of internet-based price comparisons on marketing variables, and CRM perspectives in order to achieve profitable growth of a car insurance company (Boons and Lüdeke-Freund, 2013). Girota and Netessine (2013) note that whenever a new technology emerges, there is a lack of business models to accommodate this scenario; hence this study develops an Integrated Business Re-intermediation Model (IBRM) which is potentially useful to researchers and insurance managers for enhancing the growth and profitability of insurance companies, post-price comparison.

Furthermore, the research develops a triple acquisition channel strategy framework (Figure 5.2). The framework of triple acquisition channels has extended Blattberg, *et al.* (2008)'s general model of customer choice to incorporate aggregators in the mix. This framework could have further implications for future research as it allows companies to investigate how a B2B relationship can affect a B2C relationship.

As a practical contribution, this research compares the business prospects of the case car insurance company pre- and post-joining an aggregator. The researcher has not found any research which has linked the effects of aggregators within the UK car insurance industry. The research covers aspects of car insurance business that are impacted by price comparison sites, marketing, sales and renewal rates. The research solves how the marketing mix has been impacted by aggregators, by using VAR modelling. Additionally by combining data mining and statistical models, this research is the first to combine robust regression, quantile regression, general linear models and data mining, in order to ascertain which is the most reliable in predicting customer lifetime value, for the case company. As well as solving the issue for the case company, this will enable future research and studies to consider quantile regression for predicting customer lifetime value.

8.5 Suggestions for further study

1. The emphasis of this study is to measure the effects of a UK car insurance company integrating price comparison sites into its distribution channel mix. As noted severally in the thesis chapters, the study does not consider those insurance companies that have decided not to use an aggregator. This would require additional data, which was not available for this study.

2. It would be worthwhile to gather research into the purchasing habits of customers during the car renewal period, for example whether a customer uses aggregators, would also provide some interesting findings. This will facilitate additional studies into: a) the mechanics for brand marketing when the aggregator is involved; b) wider CRM research including employee satisfaction and morale, both for aggregators and insurance companies; c) whether price aggregation cannibalises the value proposition of an insurance company; d) related issues in monitoring social networks as a customer information channel; and e) the nature of external and internal data required to support these lines of work as well as the ICT infrastructure and analytical support requirements (McCarty and Hastak, 2007), for example.

3. Aggregators are still relatively new and their full impact on CLV and retention models could not be fully explored. Further research could study whether customers who come through aggregators can be treated with the same retention program as customers who approach directly, or whether a different strategy would be more appropriate.

4. By employing a multi-channel strategy, the results generated more customers, supporting Neslin *et al.* (2006) and Blattberg *et al*. (2008). The extra customers may have been gained from lesser known brands (Leuthesser *et al.*, 1995), but this would need further research. The reason for this short coming, as mentioned previously, is due to the lack of data available for competitor car insurance companies. Another limitation to consider is the personal recommendation statistic. There is not a clear way to note if the recommendation came from friends outside or inside social networking sites.

5. Moreover, when reviewing the ambient marketing techniques, the research does not take into account any regional variations that may occur with different types of marketing being applied.

6. With particular reference to marketing insights relevant to effective business development of a typical (car) insurance company in the context of massive datasets (Big Data), this research makes connections with the following areas of work: information and data modelling (data science); insurance business development; and business modelling; strategic marketing planning, including key bottom-line metrics such as probability (ROI and CLV) and growth (customer retention). The data required for follow-on studies of the impact of aggregators on these metrics are not available in this study (see notes 2 above). Such research projects should focus on specific objectives connected with richer data on insurance claims, customer behaviours and attitudes by customer segments, linked to marketing mix variables and value propositions as highlighted by the IBRM. Hence, future work along these lines would seem to require customer survey data, external customer lifestyle data aimed at delivering research results which inform future revenue targets. Data on competitor pricing and business strategies would also be helpful. It is hoped that such augmented datasets will inform the use of more informative statistical and data mining models in the future.

References

ABI, (2012), *UK insurance key facts*, [online]. Last accessed 2nd November 2013 at: https://www.abi.org.uk/~/media/Files/Documents/Publications/Public/Migrated/Facts%20and%20figures%20data/UK%20Insurance%20Key%20Facts%202012.ashx

Ajmera, J., Ahn, H. I., Nagarajan, M., Verma, A., Contractor, D., Dill, S., & Denesuk, M. (2013, May). A crm system for social media: challenges and experiences. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 49-58). International World Wide Web Conferences Steering Committee.

Albers, S. (2012). Optimizable and implementable aggregate response modeling for marketing decision support. *International Journal of Research in Marketing*, **29** (2), 111-122.

Alford, B. and Engellhand, B. (2000). Advertised Reference Price Effects on Consumer Price Estimates, Value Perception, and Search Intention. *Journal of Business Research*, **48** (2), 93-100.

Ambler, T. (1997). How much brand equity is explained by trust? *Management Decision*, **35**(4), 283-292

Ansari A, Mela C., Nelsin S. (2008) Customer channel migration. *Journal of Marketing Research*, **45** (1), 60-76

Arbatskaya, M. (2007), Ordered search. The *RAND Journal of Economics*, **38**(1), 119–126

Ashman, S. and Clarke, K (1994), Optimising Ad Effectiveness: quantifying the major cost benefits of painstaking television pre-tests, *Admap*, **29** (2),43-56

Ataman, M., Van Heerde, H., and Mela, C. (2010). The long term Effect of marketing Strategy on Brand Sales. *Journal of Marketing Research*: **47**, (5), 866-882.

Avery, J., Steenburgh, T., Deighton, J., & Caravella, M. (2012), Adding Bricks to Clicks: Predicting the Patterns of Cross-Channel Elasticities over Time, *Journal of Marketing*, **76** (3), 96-111

Bass, F. (1969). A new product growth model for consumer durables. *Management Science*, **15**, 215-227.

Baye, M. R., Morgan J., and Scholten, P. (2004). Price Dispersion in the Small and in the Large: Evidence from an Internet Price Comparison Site. *Journal of Industrial Economics* **52** (4) pp.463-496.

BBC. (2008). *'Errors' in insurance price sites*, [online]. Last accessed 15 January 2010. Available at http://news.bbc.co.uk/1/hi/business/7389777.stm

Benoit, D. and Van den Poel, D. (2009), Benefits of quantile regression for the analysis of customer lifetime value in a contracting setting: An application in financial services. *Expert Syst. Appl.***36**, (7) 10475-10484

Blattberg, R.C. and Deighton, J. (1996) Managing marketing by the customer equity criterion. *Harvard Business Review* **74** *(*July- August), 136-144.

Blattberg, R. Kim, B. and Neslin, S. (2008), Database Marketing: Analyzing and Managing Customers (International Series in Quantitative Marketing), Publisher Springer

Blythe, J. (2005), *Principles and Practice of Marketing*, Thomson Learning

Boons, F., Leudeke-Freund, F. (2013) Business models for sustainable innovation: state of the art and steps towards a research agenda. *Journal of Cleaner Production* Vol.45, April, 9-19.

Borden, N. H. (1964). The Concept of the Marketing Mix, *Journal of Advertising Research*, **4**, (June) 2-7.

Bouwman, H., Faber, E. and Van der Spek, J. (2005). Connecting future scenario’s to business models of insurance intermediaries. Vogel, D. R., P. Walden & J. Gricar (eds), eIntegratin in Action, 18Th Bled e-Commerce Conference e Integrity. Bled Slovenia, June 6-8, 2005, pp. 1-14.

Box, GEP, Jenkins G and REissl G (2008), Time series analysis: forecasting and control 4th edition. J Wiley

Brashear,T., Kashyap, V., Musante, M., and Donthu, N. (2009) A profile of the internet shopper: evidence from six countries. *Journal of Marketing Theory and Practice,* **17**, (3), 267–281.

Breiman, L., Friedman, J., Olshen, R. and Stone, C. (1984). *Classification and regression trees*. Chapman and Hall/CRC

Broadbent, M. and Weill, P. (1997), How business and IT Managers can create IT infrastructures, *Sloan Management Review*, **38** (3), 77-92

Brown, J. and Goolsbee A., (2002), Does the internet make the markets more competitive? Evidence from the insurance industry. *Journal of Political Economy,* **110** (4), 481-507

Brynjolfsson, E. and Smith, M. (2000), 'Frictionless Commerce? A Comparison of Internet and Conventional Retailers', *Management Science*, **46**, (4), 563-85

Bucklin, G. and Gupta, S. (1999) Commercial use of UPC scanner data: Industry and academic perspectives. *Marketing Science,* **18** (3), 247-273

Casadesus-Masanell, R., & Ricart, J. E. 2010. From strategy to business models and onto tactics. *Long Range Planning*, **43**(2), 195-215.

Chan, C (2005). Online auction customer segmentation using a neural network model. *International Journal of Applied Science and Engineering*, **3** (2) .101–109.

Chen, Y., & Hu, L. (2005, August). Study on data mining application in CRM system based on insurance trade. In *Proceedings of the 7th international conference on Electronic commerce* (pp. 839-841). ACM.

Chesbrough H. and Rosenbloom, R. (2002) The role of the business model in capturing value from innovation: evidence from Xerox corporation's technology, *Industrial and Corporate Change* **11** (3), 529 - 555

Chircu, A and Kauffman R (2000), reintermediation strategies in Business-to-Business Electronic Commerce *International Journal of Electronic Commerce* **4** (4), 7-42

Clemons, E., Hitt, L. (2000). *The internet and the future of Financial Services: Transparency, Differential Pricing and Disintermediation* in R.Litan and A Rivlin(eds), *The Economic Payoff from the Internet Revolution,* Washington, DC: Brookings Institution, 87-128

Coarse , R. (1937). The Nature of the Firm, *Economica*, **4** (16), 386-404

Coelho, F. and Easingwood C. (2003), Multiple channel structures in financial services: A framework. *Journal of Financial Services Marketing* **8** (1), 22-34

Cohen, A. (2012). Asymmetric Learning in Repeated contracting: An empirical study. *Review of Economics and Statistics*, **94** (2), 419-432.

Coulter, K. and Coulter, R., (2002), Determinants of trust in a service provider: the moderating role of length relationship, *Journal of Services Marketing,* **16** (1), 35-50

Coelho, F. and Easingwood, C. (2008). A model of the antecedents of multiple channel usage. *Journal of Retailing and Consumer Services*, **15** (1), 32-41

Das, K. (2009), Relationship marketing research (1994-2006) an academic literature review and classification, *Marketing Intelligence and Planning,* **27** (3), 326-363

Datta, S., Granger,C. Barari, M. and Gibbs, T. (2007). Management of supply chain: an alternative modeling technique for forecasting. *Journal of the Operational Research Society* **58** (11), 1459-1469

David, M. (1998). *Price comparison sites: deal or no deal?* [online] last accessed 20 January 2010. Available at: http://news.bbc.co.uk/1/hi/business/7766798.stm

David Sheppard Associates, (1999), *The New Direct Marketing: How to Implement A Profit-Driven Database Marketing Strategy* 3rd edition Publisher: McGraw-Hill Professional

Dayal, S., Landesberg, H. and Zeisser, M. (2000), Building digital brands, *The McKinsey Quarterly*, **2**, 42-51.

Degraffenreid, S. (2006). Beyond advertising: why people are the new media. *Handbook of business strategy* **7** (1), 81-85

Delgado-Ballester, E. and Munuera-Alema´n, J. (2005), Does brand trust matter to brand equity? *Journal of Product & Brand Management* **14** (3), 187–196

Dekimpe, M. and Hanssens, D. (1995), The Persistence of Marketing Effects on Sales, *Marketing Science*, **14** (1), 1–21.

Dekimpe, M. and Hanssens, D. (1999), Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability, *Journal of Marketing Research*, **36** (November), 397–412.

Dekimpe, M. and Hanssens D. (2000), Time series models in marketing; Past, present and future *International Journal of research Marketing* **17** (2),183-193

Delgado-Ballester, E. and Munuera-Alemán, J., (2005), Does brand trust matter to brand equity? *Journal of Product & Brand Management*, **14** (3), 187 - 196

Dictionary.com, "disintermediation," in *The American Heritage® Dictionary of the English Language, Fourth Edition*. Source location: Houghton Mifflin Company, 2004.

[online] last accessed January 21 2010. Available at: http://dictionary.reference.com/browse/disintermediation.

Doherty, N. and Ellis-Chadwick, F (2010). Internet retailing: the past, the present and the future. *International Journal of Retail & Distribution Management*, **38**. (11/12), 943–965.

Domowitz, I. (2002). Liquidity, transaction costs, and reintermediation in electronic markets. *Journal of Financial Services Research,* **22** 1(/2), 141-157

Donkers, B., Verhoef, P. and de Jong, M. (2007). Modelling CLV: A Test of Competing Models in the Insurance Industry, *Quantitative Marketing and Economics,* **5** (2), 163–190.

Doyle, P. (1989), Building successful brands: the strategic objectives, *Journal of Marketing Management,* **5** (1), 77-95.

Drew, J., Mani, D., Betz, A. and Datta, P. (2001), Targeting customers with statistical and data mining techniques, *Journal of Service Research,* **3** (3), 205-219

Drucker, P. (2007), *The practice of management (classic drucker collection)*, 2nd Edition. Publisher: A Butterworth-Heinemann Title.

Duboff, R. (1992). Marketing to maximise profitability. *The journal of business strategy*, **13** (6), 10-13

Duffy, R., Fearne, A., Hornibrook, S., Hutchinson, K., & Reid, A. (2013). Engaging suppliers in CRM: The role of justice in buyer–supplier relationships. *International Journal of Information Management*, ***33***(1), 20-27.

Dumm, R.and Hoyt, R..(2003). Insurance distribution channels: Markets in transition. *Journal of Insurance Regulation*, **22**(1), pp27-47.

E-banking, nodate, *glossary*, [online] last accessed 17 November 2008. Available at: http://www.ffiec.gov/ffiecinfobase/booklets/e_banking/ebanking_04_appx_b_glossary.html

Enders, Walter (1995), *Applied Econometric Time Series*. New York: John Wiley & Sons.

Engle, R. and Kroner, K., (1995). Multivariate simultaneous generalized ARCH. *Econom Theory* **11** (1), pp.122–150.

Ellison, G and Ellison S. (2009), Search, Obfuscation, and Price Elasticities on the Internet, *Econometrica*, **77** (2), 427-452.

Engle, R., 2001, GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics, *Journal of Economic Perspectives*, **15** (4), 157–168

Evans, M., (2002), Prevention is better than cure: Redoubling the focus on customer retention, *Journal of Financial Services Marketing*, **7** (2), 186-198.

Even, A.S., Shankaranarayan, G. and Berger, P. (2010) Inequality in the utility of customer data: Implications for data management and usage. Journal of Database Marketing & Customer Strategy Management **17**(1), 19–35.

Foss, B., Stone, M., & Ekinci, Y. (2008). What makes for CRM system success—Or failure? *Journal of Database Marketing & Customer Strategy Management*, ***15***(2), 68-78.

Gamper, H. C. (2012). How Can Internet Comparison Sites Work Optimally for Consumers? *Journal of Consumer Policy*, **35** (3), 333-353.

Gee, R., Coates, G. and Nicholson, M. (2008) Understanding and profitably managing customer loyalty. *Marketing Intelligence & Planning* **26** (4), 359-374

Gellman, R. (1996). Disintermediation and the Internet, *Government Information Quarterly* **13**, (1), 1-8

Ghosh, M. and John, G. (2000). Experimental Evidence for Agency Models of Salesforce Compensation, *Marketing Science*, **19**, (4), 348-365

Ghosh, S (2013), BBC admits £100 million IT project was a waste [online], last accessed 23 November 2013. Available at: http://www.pcpro.co.uk/news/enterprise/382057/bbc-admits-100-million-it-project-was-a-waste

Giaglis, G. M., Klein, S., and O'Keefe, R. M. (1999). Disintermediation, reintermediation, or cybermediation? The future of intermediaries in electronic marketplaces. In S.Klein *et al*. (Eds.), Twelfth Bled Electronic Commerce Conference on "Global Networked *Organizations"* (pp. 389–407). Bled (Slovenia).

Giaglis, G. (2002). *Towards a predictive model for electronic marketplaces intermediation opportunities* [online]. Last Accessed 23 March 2009. Available at: http://www.mobiforum.org/mobiforum/WRC/Papers/wcmec02_Giaglis_marketplaces.pdf

Giaglis, G. M., Klein, S., and O'Keefe, R. M (2002), The role of intermediaries in electronic marketplaces: developing a contingency model. *Information Systems Journal* **12** (3), 231–246.

Girotra, K. and Netessine, S. (2013), Business Model Innovation for Sustainability *Working paper*
http://www.insead.edu/facultyresearch/research/doc.cfm?did=52400

Goodhue, D., Wixom, B. and Watson, H. (2002). Realizing business benefits through CRM: hitting the right target in the right way. *MIS Quarterly Executive*, **1** (2), 79–94.

Godes, D., Mayzlin, D. (2004). Using Online Conversations to Study Word-of-Mouth Communication, *Marketing Science*, **23** (4), 545-560

Gorman, M. Brannon, I, Salisbury, W. (2009). Who wins when price information is more ubiquitous? An experiment to assess price comparison services and how they influence prices, **19** (2/3), 151-162

Granados, N. F., Gupta, A., and Kauffman, R. J. (2006). The Impact of IT on Market Information and Transparency: A Unified Theoretical Framework, *Journal of the Association for Information Systems*, **7** (3), 148-178.

Granados, N. F., Kauffman, R. And King, B. (2008), the emerging role of vertical search engines in travel distribution: a newly-vulnerable electronic markets perspective. *Proceeding of the 41st Hawaii International Conference in System Sciences*

Granados, N., Gupta, A, and Kauffman, R.J. (2011). Online and Offline Demand and Price Elasticities: Evidence from the Air Travel Industry, *Information Systems Research*, **23** (1), 164-181

Granger, C., (1969), Investigating Causal Relations by Econometric Models and Cross-Spectral Methods, *Econometrica*, **37** (3), 424–38.

Grewal, D, Monroe, K. and Krishnan R (1998), The effects of price-comparison advertising on buyers perceptions of acquisition value, transaction value and Behavioral Intentions, *Journal of Marketing,* **62** (2), 46-59

Gronhaug, K. and Gilly, M. (1991), A transactional cost approach to consumer dissatisfaction and complaint actions, *Journal of Economic Psychology*, **12** (1), 165-183

Guelman, L., Guillén, M., & Pérez-Marín, A. M. (2012). Random Forests for Uplift Modeling: An Insurance Customer Retention Case. *Modeling and Simulation in Engineering, Economics and Management. Lecture Notes in Business Information Processing.* **115** (2012), 123-133

Guillén, M., Pérez Marín, A. M., & Alcañiz, M. (2011). A logistic regression approach to estimating customer profit loss due to lapses in insurance. *Insurance markets and companies: analysis and actuarial computations* **2**, (2), 42-54

Haaker, T., FaberE. and Bouwman H., (2006), Balancing customer and network value in business models for mobile services, *International Journal of Mobile Communications*, **4** (6), 645-661

Hamel, G., (2002) *Leading the Revolution*, Publisher: New American Library; Reprint edition

Hammonds, K. (2003), How Google grows… and grows… and grows…, fast company, *[online]* last accessed 10 August 2012. Available at: *http://www.fastcompany.com/46495/how-google-growsand-growsand-grows*

Hampel, F. R. (1978). Modern trends in the theory of Robustness. *Statistics*, **9** (3), 425-442

Han, J., Kamber, M. and Pei, J. (2011). *Data Mining: Concepts and Techniques, 3rd edition* Morgan Kaufmann Publishers

Haan, M. A., & Moraga-González, J. L. (2011). Advertising for Attention in a Consumer Search Model*. *The Economic Journal*, **121** (552), 552-579.

Hanssens, D., Parsons, L. and Schultz, R. (2003). *Market Response Models: Econometric and Time Series Analysis.* 2nd ed. Edition, Springer

Hanssens, D., Leeflang, P. and Wittink, D. (2005). Market response models and marketing practice. *Appl. Stochastic Models Bus. Ind.* **21**(4-5), 423–434

Hawkes VA (2000). The Heart of the Matter: The Challenge of Customer Lifetime Value. *Customer Relationship Management Forum Resources,* (Vol.13 pp.1-10)

Harris, L. and Coles, A. (2004). The impact of the Internet on marketing priorities in the financial services industry, *Strategic Change* **13**(1),11–18

Harrison, T., (2000) *Financial Services Marketing* Publisher : Pearson Education Ltd.

Hiekkanen, K., Helenius, M., Korhonen, J. J., & Patricio, E. (2013). Aligning Alignment with Strategic Context: A Literature Review. *Digital Enterprise Design and Management* **205,** 81-98

Hill, C. R., Griffiths, W.E. and Lim, G.C. (2012). Principles of econometrics, John Wiley & Sons, Inc.

Hoekstra, J. and Huizingh, K. (1999), The Lifetime Value Concept in Customer Based Marketing, *Journal of Market Focused Management* **3** (3/4), 257-274

Holland, C. P., & Mandry, G. D. (2013, January). Online Search and Buying Behaviour in Consumer Markets. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (pp. 2918-2927). IEEE.

Hoyt, R., Dumm, R. and Carson, J. (2006). An Examination of the Role of Insurance Producers and Compensation in the Insurance Industry. *[online],* last accessed 12 October 2009, Available at:
http://reference.kfupm.edu.sa/content/e/x/executive_summary_an_examination_of_the__81190.pdf

Huang, W., Swaminathan, J. (2009). Introduction of a second channel: implications for pricing and profits. *European Journal of Operational Research* **194 (**1), 258–279.

Huang, R., & Sarigöllü, E. (2012). How brand awareness relates to market outcome, brand equity, and the marketing mix. *Journal of Business Research*, **65** (1), 92-99.

Hwang, H. Jung, T. and Suh, E. (2004), An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry, *Expert Systems with Applications* **26**(2).181–188

Institute of Insurance broker. (2003). Direct Insurers, *[online]*. Last accessed 19 December 2008. Available at:
https://vault2.secured-url.com/iib/encyclopeadia_item.asp?encyclopedia_id=151

Islam, S. M. N. and Watanapalachaikul, S. (2005). *Empirical finance: Modelling and analysis of emerging financial and stock markets.* New Work, Springer.

Jallat, F. and Capek, M. (2001). Disintermediation in Question: new economy, new networks, new middlemen. *Business Horizons*, **44**(2), 55-60

Jandos, J. (2001). On the role of intermediaries in digital economy. *System integration* pp.503-508

Johansen, S. (1995), A Statistical Analysis of Cointegration for I(2) Variables, *Econometric Theory*, **11**(01), pp.25–59.

Johansen, S. (1988), Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control, Vol.12, pp.231–254.

Johansen, S. and Juselius, K. (1990), Maximum Likelihood Estimation and Inference on Cointegration: With Applications to the Demand for Money, Oxford Bulletin of Economics and Statistics, **52**(2), 169–210.

Johnson E., Hershey, J., Meszaros, J. and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, **7**(1), 35-51

Johnson, R. and Onwuegbuzie, A. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, **33**, (7), 14-26

Joshi, S. and Hanssens D. (2009). Movie Advertising and the Stock Market Valuations of Studios: A case of "Great Expectations?, *Marketing Science,* **28**, (2), 239-250

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics,* **29** (2),119-127

Kauffman, (2013).Digital Intermediation in Electronic Commerce . the eBay Model,. *Electronic Commerce and Virtual Business*, Stuart Barnes and Brian Hunt (eds.), Butterworth-Heinemann, Oxford, pp. 45-66.

Keller, K. (1993), Conceptualizing, measuring and managing customer-based brand equity, **57** (1), 1-22

Keller, K. and Lehmann, D. (2006) Brands and Branding: Research Findings and Future Priorities *Marketing Science* **25** (6), 740–759.

Ketchen, D. and Hult, T. (2007). Bridging organization theory and supply chain management: The case of best value supply chains. *Journal of Operations Management,* **25** (2),573-580

Kim, J. Suh, E. And Hwang, H. (2003), A model for evaluating the effectiveness of CRM using the balanced scorecard, *Journal of interactive marketing*, **17** (2), 5-19

Kim, S., Jung, T., Suh, E. and Hwang, H. (2006). Customer segmentation and strategy development based on customer lifetime value: A case study. *Expert Systems with Applications* **31** (1), 101–107

Kim, H-S., Kim, Y-G. (2009) A CRM Performance Measurement Framework: Its Development Process and Application, *Industrial Marketing Management*, **38** (4), 477-489

Knox, S. (2004), Positioning and branding your organisatio.n *Journal of Product and brand management* **13** (2),105-115

Koenker, R. and Gilbert, B. (1978). Regression Quantiles. *Econometrica.* **46** (1), 33–50.

Kotler, P. (1997). *Marketing management: Analysis, Planning, Implementation and Control*, 8th Edition Prentice Hall College Division

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of econometrics*, ***54***(1), 159-178.

Labusch, N., & Winter, R. (2012). Method Support of Large-Scale Transformation in the Insurance Sector: Exploring Foundations. In *Trends in Enterprise Architecture Research and Practice-Driven Research on Enterprise* Transformation **131** pp. 60-78

Laffey, D. (2009). Click trading: A case study of Moneynet, *Journal of Strategic Information Systems* **18**(1), 56–64

Laffey, D. (2010), Comparison websites: evidence from the service sector. *The service Industries Journal*, **30** (12), 1939-1954

Lages, L. F., Lancastre, A., & Lages, C. (2008). The B2B-RELPERF scale and scorecard: Bringing relationship marketing theory into business-to-business practice. *Industrial Marketing Management*, ***37***(6), 686-697.

Lam , S., Shankar, V. Erramilli, M. and Murthy B. (2004), Customer value, satisfaction, loyalty and switching costs: an illustration from a business-to-business service context, *Journal of Marketing Science*, **32** (3), 293-311

Lambert, S. C., & Davidson, R. A. (2013). Applications of the business model in studies of enterprise success, innovation and classification: an analysis of empirical research from 1996 to 2010. *European Management Journal*, ***31***(6), 668-681.

Lampe, C., Ellison, N., and Steinfeld, C. (2007). A familiar Face(book): Profile elements as signals in an online social network. *Proceedings of Conference on Human Factors in Computing Systems* (pp. 435–444). New York: ACM Press.

Lee, J., Park, D. and Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, **7** (3), 341-352.

Leeflang, P., Wittink, M, and Philippe N. (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic Publishers.

Leuthesser, L., Kohli. C. and Harich, K. (1995), Band equity: the halo effect measure, *European Journal of Marketing*, **29** (4), 57-66

Li, H, Kuo, C, Rusell, M., (1999), The Impact of Perceived Channel Utilities, Shopping Orientations, and Demographics on the Consumer's Online Buying Behavior *Journal of Computer-Mediated Communication*, **5** (2) (online)
Last accessed 11 April 2011
http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1999.tb00336.x/full

Libai, B., Bolton, R., Büegel, M. S., de Ruyter, K., Götz, O., & Risselada, H. (2012). Decomposing the Value of Word of Mouth Seeding Programs: Acceleration vs. Expansion. *Journal of Marketing Research*. **77** (3), 1-14.

Linde, F. (2009) Pricing information goods, *Journal of Product & Brand Management*, **18** (5), 379 - 384

Lutkepohl, H and Hans-Eggert, R. (1992). Impulse- Response Analysis of Cointegrated Systems, *Journal of Economic Dynamics and Control*, **16** (1), 53-78.

Malhotra, N. K., Birks, D. F., Wills, P. 2012. *Marketing Research. An Applied Approach*. Prentice Hall. 4th Edition.

Malthouse, E. and Blattberg, R., (2005). Can we predict customer lifetime value? *Journal of interactive marketing*. **19** (1), 2 – 16

Mata, F., Fuerst, W. and Barney, J. (1995) Information technology and sustained competitive advantage: a resource-based analysis, *MIS Quarterly*, **19** (4), 487-505,

McDonald, M, Chernatony, L. ,Harris, F. (2001), Corporate marketing and service brands Moving beyond the fast-moving consumer goods model, *European Journal of Marketing*, **35** (3/4), 335-352.

McDonald, S. and Wren, C. (2009), ‘Price Dispersion and the Ability to Search: The UK Internet Motor Insurance Market’, *Newcastle Discussion Papers in Economics*, no. 2009 / 07, 1361 - 1837

McDonald, S., & Wren, C. (2012). Informative brand advertising and pricing strategies in internet markets with heterogeneous consumer search. *International Journal of the Economics of Business,* **19** (1), 103-117.

Meyer, C. and Schwager, A. (2007). Understanding customer experience. *Harvard Business Review,* **85** (2), 116–126.

Mizaei, T. and Iyer, L. (2014, March) Predictive Analytics in CRM: Literature Review. In *Proceedings of the Southern Association for Information Systems Conference, Macon,* GA, USA March 21st–22nd, 2014

Morgan, J., Orzen, H. and Sefton, M., (2006) An experimental study of price dispersion. *Games and Economic Behaviour* **54** (1), 134–158

Mukerjee, K. (2013). Strategizing for CRM to leverage its benefits. *Business Strategy Series*, ***14***(4), 118-122.

Müller, C. and Zimmermann, H (2003) Beyond mobile: research topics for upcoming technologies in the insurance industry. *Paper presented at the Proceeding of the 36th Hawaii International Conference on System Sciences*, HI, USA, 5–9 January.

Murray, M. (1994), A drunk and her dos: an illustration of cointegration and error correction. *The American Statistician* **48** (1), 37-39

Murphy, J. (1992). What is Branding? (Branding a Key Marketing Tool), Macmillan, Basingstoke.

Murphy, J., Burton, J., Gleaves, R. and Kitschoff, J. (2006), *Converting Customer Value from retention to profit*. Jon Wiley & Sons

Neelamegham, R. and Chintagunta, P. (1999). A Bayesian Model to Forecast New Product Performance in Domestic and International Markets. *Marketing Science,* **18** (2), 115-136.

Neslin, S., Grewal, D., Leghorn, R., Shankar, V., Teerling, M., Thomas, J. and Verhoef, P., (2006), Challenges and opportunities in multi-channel customer management, *Journal of Service Research*, **9** (2), 95-112

Ng, I., (2004), Does direct marketing need to have a direction? *Marketing Intelligence & Planning*, **23** (7), 628-635

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, **36**(2), 2592-2602.

Ofcom (2011) *International Communications Market Report 2011* [Online]. Last accessed 1 April 2014. Available at: http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr11/icmr/ICMR2011.pdf

Oliveira, A., Sampaio do Prado, J. , Cysneiros, L, Cappelli, C. (2007). *Eliciting Multi-Agent Systems Intentionality: from Language Extended Lexicon to i* Models*, sccc, pp.40-49, XXVI International Conference of the Chilean Society of Computer Science (SCCC'07)

Olson, D., Cao, Q., Gu, C. and Lee, D. (2009). Comparison of customer response models *Service Business* **3** (2), 117-130.

Osterwalder, A. and Pigneur, Y., (2002) An e-Business Model Ontology for Modeling e-Business," *Proceedings of the 15th Bled Electronic Commerce Conference*, 1–12, Bled, Slovenia (June 2002).

Pantelidaki, S. and Bunn, D. (2005). Development of a Multifunctional Sales Response Model with the Diagnostic Aid of Artificial Neural Networks. *Journal of Forecasting* **24** (7), 505–521

Papadopoulos, S. (1987). Strategic Marketing techniques in international tourism. *International Marketing Review* **4** (2), 71-84

Papatla, P., Liu and F. (2009). Google or BizRate? How search engines and comparison sites affect unplanned choices of online retailers. *Journal of Business Research* **62** (11), 1039–1045

Pauwels, K. (2004). How Dynamic Consumer Response, Competitor Response, Company Support, and Company Inertia Shape Long-Term Marketing Effectiveness, *Marketing Science* **23** (4), 596-610

Pauwels, K. and Hanssens, D. (2007). Performance Regimes and Marketing Policy Shifts. *Marketing Science*, **26** (3), 293-311

Pauwels, K and Neslin, S. (2008), Building with Bricks and Mortar:The Revenue Impact of Opening Physical Stores in a Multichannel Environment, *Working Paper*, Hanover, NH: Tuck School of Business, Dartmouth College.

Payne, A., & Frow, P. (2006). Customer relationship management: from strategy to implementation. *Journal of Marketing Management*, ***22***(1-2), 135-168.

Peppers, D. and Rogers, M. (1998) *The One to One Future: Building Relationships One Customer at a Time.* Publisher: Bantam Doubleday Dell Publishing Group

Payne, A..and Frow, P.,(2005). A strategic framework for customer relationship management, *Journal of marketing,* **69** (4), 167-176

Peterson, R., Balasubramanian, S., and Bronnenberg, B. (1997). Exploring the implications of the Internet for consumer marketing. *Journal of the Academy of Marketing Science*, **25** (4), 329–346.

Petrovic, O., Kittl, C. and Tekstan, R. (2001). Developing Business Models for eBusiness. *Proceedings of the International Conference on Electronic Commerce*

Pfeifer, P. (2005), The Optimal Ratio of Acquisition and Retention Costs, *Journal of Targeting, Measurement, and Analysis for Marketing*, **13**, (2), 179-188.

Pfeifer, P., Haskins, M., Conroy, R., 2005. Customer lifetime value, customer profitability, and the treatment of acquisition spending. *Journal of Managerial Issues* **17** (1), 11-25.

Pfeil, C., Posselt, T. and Maschke, N. (2008) Incentives for sales agents after the advent of the internet, *Marketing Letters*, **19** (1), 51-63.

Porter, M. (1998), Competitive Strategy: Techniques for Analyzing Industries and Competitors, Publisher: Simon & Schuster Ltd; New edition

Porter, M (2001), Strategy and the internet, *Harvard Business Review march*, **79** (3), 63-78

Porter, M., (2004), Competitive Advantage : Free Press; New edition (19 Jan 2004)

Rangaswamy, A. and Bruggen, G., (2005). Opportunities and challenges in multichannel marketing: An introduction to the special issue. *Journal of Interactive Marketing*, **19** (2), 5-11

Reichheld, F and Sasser W., (1990), Zero Defections: quality comes to services, *Harvard Business Review*, **68 (**September-October), 105-111

Reichheld, Frederick F. (1993), Loyalty-Based Management, *Harvard Business Review*, **71** (2), 64-73.

Reichheld, Frederick E. (1996). Learning From Customer Defections. *Harvard Business Review* **74** (March-April): 56-69.

Reicheld, F. (2001). *The Loyalty Effect: The Hidden Force Behind Growth, Profits, and Lasting Value [Paperback]* Harvard Business School Press

Reichheld, F., (2006). *The ultimate question driving good profits and true growth.* Publisher Harvard Business School Publishing Corporation

Reuters, 2010. *Nokia says free satnav to lift services sales,* [online]. Last accessed 14 February 2010. Accessed at : http://www.reuters.com/article/idUSTRE61E4EF20100215?type=technologyNews

Reuver M., Haaker, T. (2009), Designing viable business models for context-aware mobile services *Telematics and Informatics*, **26** (3), 240-248

Rios, R. and Riquelme, H. (2008). Brand equity for online companies. M*arketing Intelligence & Planning,* **26** (7), 719-742

Robertshaw, G. (2011). An examination of the profitability of customers acquired through price comparison sites: Implications for the UK insurance industry. *Journal of Direct, Data and Digital Marketing Practice* **12** (3), 216–229

Rose, J. (2008). *Overview of the UK insurance market during 2007,* [online]. Last accessed 14 November 2008. Available at: http://www.abi.org.uk/Display/File/524/Annual_Overview_2007.pdf

Rosenberg, L.G. and Czepiel, J.A. (1984). A marketing approach for customer retention. *Journal of Consumer Marketing,* **1** (2), pp.45- 51.

Rousseeuw, P. and Leroy, A. (2003). *Robust Regression and Outlier Detection,* Wiley-Blackwell

Rowley, J., (2004). Just another channel? Marketing communications in e-business. *Marketing Intelligence & pllanning,* **22** (1), 24-41.

Rust, R. Ambler,T., Carpenter, G. Kumar, V. and. Srivastava, R. (2004). Measuring Marketing Productivity: Current Knowledge and Future Directions. *Journal of Marketing* **68** (4), 76-89

Rust, R. T. and Verhoef, P. C. (2005). Optimizing the marketing mix interventions in intermediate-term CRM. *Marketing Science*, **24**(3), 477–489

Ryals, L. and Knox S. (2005). Measuring Risk-adjusted Customer Lifetime Value and its Impact on Relationship Marketing Strategies. *European Journal of Marketing*, **39** (5/6), 456-472

Rygielski C., Wang, J., Yen, D., 2002, Data mining techniques for customer relationship management, *Technology in Society* **24** (4), 483–502

Samaniego, M., Arranz, A. and Cabezudo, R. (2006). Determinants of internet use in the purchasing process *Journal of Business & Industrial Marketing* **21** (3), 164–174.

SAS. 2003. Data mining using SAS enterprise miner: a case study approach: 2nd edition Publisher SAS institute,

SAS, 2008, SAS/ETS 9.2 User's Guide: Publisher: SAS Publishing

Sethuraman, R., Srinivasan, V. and Doyle, K. (1999), Asymmetric and Neighborhood Cross-Price Effects: Some Empirical Generalizations, *Marketing Science*, **18** (1), 23-41.

Shapiro, C. (1983). Premiums for High Quality products as Returns to Reputations. *Quarterly Journal of Economics*, **98** (4), 659-679.

Shafer, S., Smith, H. and Linder, J. (2005), The power of business models. *Business Horizons* **48**(3), 199-207

Sharp B, Wright M., and Goodhardt G. (2002) Purchase loyalty is polarised into either repertoire or subscription patterns. *Australasian Marketing Journal* **10** (3), 7 – 20.

Shen, C. and Chuang, H. (2009), A Study on the Applications of Data Mining Techniques to Enhance Customer Lifetime Value. *WSEAS transactions on information Science and applications* **6**, (2), 319-328

Shih, B. Y., Chen, C. Y., & Chen, Z. S. (2012). An empirical study of an internet marketing strategy for search engine optimization. *Human Factors and Ergonomics in Manufacturing & Service Industries*. **23** (6), 528–540

Shon, T., Parker, C. and Swatman, P. (2003). *Characteristics of Australian B2B iMarketplaces. [online]* Last accessed 19 March 2009.
Available at: http://www.pacis-net.org/file/2003/papers/market-spaces/92.pdf

Simon, C. and Sullivan, M. (1993). The Measurement and Determinants of Brand Equity: A Financial Approach. *Marketing Science,* **12** (1), 28-52.

Simon, E. (2011), Comparison sites explained [online]. Last accessed 8 March 2012. Available at:
http://www.telegraph.co.uk/finance/personalfinance/consumertips/8368276/Comparison-sites-explained.html#

Sims, C (1980), Macroeconomics and Reality, *Econometrica*, **48** (1), 1–48.

Smith, J. (2008). *Qualative psychology: a practical guide to research methods,* 2nd edition. Sage Publications Inc.

Smith K., Willis, R. and Brooks, M., (2000). An analysis of customer retention and insurance claim patterns using data mining: A case study. *The Journal of the Operational Research Society*, **51**, (5), 532- 541

Sohrabi B, Khanlari A (2007). Customer lifetime value (CLV) measurement based on RFM model. *Iranian Acc. Aud. Rev*., **14** (47), 7-20.

Sophonthummapharn, K. (2009). The adoption of techno-relationship innovations, A framework for electronic customer relationship management, *Marketing Intelligence & Planning,* **27** (3), 380-412

Srinivasan, K., Kekre, S. and Mukhopadhyay, T. (1994), Impact of Electronic Data Interchange Technology on JIT Shipments. *Management science* **40** (10), 1291-1304

Srinivasan, S and Dominique H (2007), Marketing and Firm Value: Metrics, Methods, Findings and Future Directions, working paper, Andersons School of Management,University of California, Los Angeles

Stern, J. and Priore, A. (2000) Email marketing: using email to reach your target Audience and build customer relationships. *John Wiley & Sons New York.*

Stigler, G. (1961). The economics of information, *The Journal of Political Economy,* **69** (3), 213-225

Stone, M. and Foss, B. (2002). *CRM in Financial Services: A Practical Guide to Making Customer Relationship Management Work*
Kogan Page Ltd

Strauss, A and Corbin, J. (1998). *Basics of Qualitative research, techniques and procedures for Developing Grounded Theory,* 2nd edition
Sage Publications, Inc.

Sweeney, J. and Swait, J. (2008). The effects of brand credibility on customer loyalty. *Journal of Retailing and Consumer Services,* **15** (3),179-193

Syal, R. (2013) Abondoned NHS IT system has cost £10bn so far [online]. Last accessed 23 November 2013, available at:
http://www.theguardian.com/society/2013/sep/18/nhs-records-system-10bn

Tapscott, D. (2009) *Grown Up Digital: How the net generation is changing your world*
Publisher: McGraw-Hill Professional

Taylor, S., (2001), Assessing the use of regression analysis in examining service recovery in the insurance industry: relating service quality, customer satisfaction and customer trust. *Journal of Insurance Issues*, **24** (1/2), 30–57.

Teece, R. (2010), Business Models, Business Strategy and Innovation *Long Range Planning* **43** (2), 172-194

Telang, P., and Singh, M. (2009), Business Modeling via commitments Service-Oriented Computing: Agents, Semantics, and Engineering Lecture Notes in Computer Science, 2009, Volume 5907/2009, pp.111-125, DOI: 10.1007/978-3-642-10739-9_9

Thomas, G. 2012, Non-Risk Price Discrimination in Insurance: Market Outcomes and Public Policy, *The Geneva Papers*, **37** (1), 27–46

Thomas, J., Reinartz, W. and Kumar, V. (2004), Getting the most out of all your customers, *Harvard Business Review*, **82** (7-8), 116-23.

Thorne, A. (1971). The marketing mix- a problem analysis. *European Journal of Marketing*, **5** (4), 207-216

Trachtenberg, J. and Fowler, G. (2010). *Amazon, Macmillan Settle Price Dispute* [online]. Last accessed 15 February 2010. Available at: http://online.wsj.com/article/SB10001424052748704197104575051553263647896.html?mod=WSJ_hpp_sections_business

Thomas, R. G. (2012). Non-risk price discrimination in insurance: market outcomes and public policy. *The Geneva Papers on Risk and Insurance-Issues and Practice*, **37** (1), 27-46.

Thuring, F., Nielsen, J. P., Guillén, M., & Bolancé, C. (2012). Selecting prospects for cross-selling financial products using multivariate credibility. *Expert systems with Applications*, **39** (10), 8809-8816.

Trusov, M., Bucklin, R. & Pauwels, K. (2009), Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing* **73** (5), 90–102

UK Statue Law database, no date, *Road Traffic Act 1988 (c. 52), [online]*. Last accessed 14 November 2008. Available at: http://www.statutelaw.gov.uk/content.aspx?LegType=All+Legislation&title=Road+Traffic+Act+1988&searchEnacted=0&extentMatchOnly=0&confersPower=0&blanketAmendment=0&sortAlpha=0&TYPE=QS&PageNumber=1&NavFrom=0&parentActiveTextDocId=2276534&ActiveTextDocId=2276785&filesize=99497

Verhoef, P. and Donkers, B. (2001). Predicting customer potential value an application in the insurance industry. *Decision Support Systems,* **32** (2), 189–199

Verhoef, P. and Donkers, B. (2005). The effect of Acquisition Channels on Customer Loyalty and Cross Buying. *Journal of Interactive Marketing*, **19** (2), 31-43

Verhoef, P.C., Venkatesan, R., McAlister, L., Malthouse, E.C., Krafft, M. and Ganesan, S. (2010) CRM in data-rich multichannel retailing environments: A review and future research directions. Journal of Interactive Marketing **24**(2): 121–137.

Vries, G, (2012), Online display ads: the brand awareness black hole *[online]*. Last accessed 26 July 2013. Available at: http://www.forbes.com/sites/ciocentral/2012/05/07/online-display-ads-the-brand-awareness-black-hole/

Waldfogel, J. and Chen, L. (2006), Does information undermine brand? Information intermediary use and preference for branded web retailers. *The journal of industrial economics*. **54** (4), 425-449

Wang, K., Yang, Q and Yeung J., (2005) Mining Customer Value: From Association Rules to Direct Marketing *Data Mining and Knowledge Discovery*, 11 (1), 57–79

Webb, K (2002). Managing channels of distribution in the age of electronic commerce. *Industrial marketing Management*, **31** (2), 95-102

Wei, J.-T., Lin, S.-Y. and Wu, H.-H. (2010). The review of the application of RFM model, *African Journal of Business Management*, **4** (19), 4199-4260.

Weltevreden, J. (2007), Substitution or Complementarity? How the Internet Changes City Centre Shopping, *Journal of Retailing and consumer Services*, **14** (3), 192-207.

Wilson, R., Gilligan, C. (2008), *Strategic marketing management: planning implementation and control*, 3rd Edition, publisher: Elsevier

Wolk, A. and Skiera, B. (2009). Antecedents and consequences of Internet channel performance. *Journal of Ret\ailing and Consumer Services*, **16** (3), 163-173

Wu, B. Ye, Q. and Yang, S. (2009) Group CRM: a new telecom framework from social network prospective. *Proceedings of the 1st ACM international workshop on Complex networks meet information & knowledge management*, November pp.3-10

Villanueva, J, Yoo, J. and Hanssens, D. (2008), The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth, *Journal of Marketing Research*, **45** (1), 48–59.

Yoo, B. and Donthu, N. (2001), Developing and validating a multidimensional consumer-based brand equity scale, *Journal of Business Research*, **52** (1), 1-14.

Yoo, S., and Hanssens D. (2005), Modeling the Sales and Customer Equity Effects of the Marketing Mix, working paper, University of California, Los Angeles, Anderson School of Management. 2

Yoo, S. & Hanssens, D.M. (2008). Measuring marketing effects on customer equity for frequently purchased brands, *UCLA Marketing Studies Center Working paper*, *April*. http://www.anderson.ucla.edu/faculty/dominique.hanssens/content/CESP_12APR08%20(2).pdf

Zahay, D., Peltier, J. and Krishen, A.S. (2012) Building the foundation for customer data quality in CRM system quality for financial services firms. *Journal of Database Marketing & Customer Strategy Management* **19**(1): 5–16.

Zott, C., Amit, R., & Massa, L. (2011). The business model: recent developments and future research. *Journal of management*, ***37***(4), 1019-1042.

Appendix

Appendix 1.1: Brief background on the UK car insurance industry and price comparison sites

1.1 Introduction

Car insurance is a legal requirement within the UK, for all car owners, even if the owner of the vehicle does not use their vehicle. There are three main type of motor insurance:

- Third party - this is the minimum cover required by law. This will only cover costs to the other party in an event of an accident resulting in injuries or damage to their vehicle.
- Third part fire and theft - as above but with added protection for the vehicle in circumstance of the vehicle being stolen or destroyed in a fire.
- Comprehensive - the highest level of cover. As above, but this also covers the insurers damage to their car in a result of car accident.

With different aspects of insurance cover available along with the choice of numerous insurance companies available, this has led to high switching costs between insurance companies.

This section describes the impact of price comparison sites within the UK car insurance industry with the following objectives:

- To discuss the intermediation, disintermediation and reintermediation
- To provide an overview of the strategic marketing within the car insurance industry
- To present a background to customer segmentation for targeted customer retention

The background information provides key insights into the UK car insurance industry pre- and post-comparison sites. As mentioned previously, this understanding is significant to explaining the consequences of the research findings.

1.2. Intermediation, disintermediation and reintermediation

Traditionally people in the UK bought car insurance through brokers. These were intermediaries, which provided a service for the customers by acting as a go-between between companies and the customer. The arrival of direct insurers has completely changed the way people buy insurance.

Direct insurers ‘were born from an original concept of selling direct to the public, via the telephone in the early 1980s, when Direct Line (providing motor insurance) was launched’ (*Institute of Insurance Brokers*, 2003). This act of cutting out ‘the middle man’ (intermediary) is usually referred to as disintermediation - defined as the ‘withdrawal of funds from intermediary financial institutions, such as banks and savings and loan associations, in order to invest in instruments yielding a higher return' (*American Heritage Dictionary*, 2004). The term was coined in the late 1960s to describe the process in which investors realised that ‘banks were no longer needed to serve as exclusive intermediaries between small depositors and the financial markets’ (Gellman, 1996, p.2).

Although the selling of insurance over the phone was the catalyst for disintermediation, the advent of the internet has allowed this process to progress even further, with some companies only conducting their business online. ‘Net-based direct interaction eliminates the role previously enjoyed by financial advisors, retail stock brokers, and insurance agents’ (Clemons and Hitt 2000, pp.5).

While disintermediation may have closed down some avenues of business due to the electronic market places making the value chains shorter (Giaglis *et al*. 1999), Jallet and Capek (2001, pp.60) view it as a ‘means [of] discovering new distribution models’. Disintermediation has re-introduced intermediation in the form of

- *Reintermediation – traditional intermediaries may find opportunities to leverage their expertise and economies of scale*
- *Cybermediation scenario- the advent of electronic markets will create unprecedented opportunity for wholly new types of intermediaries*

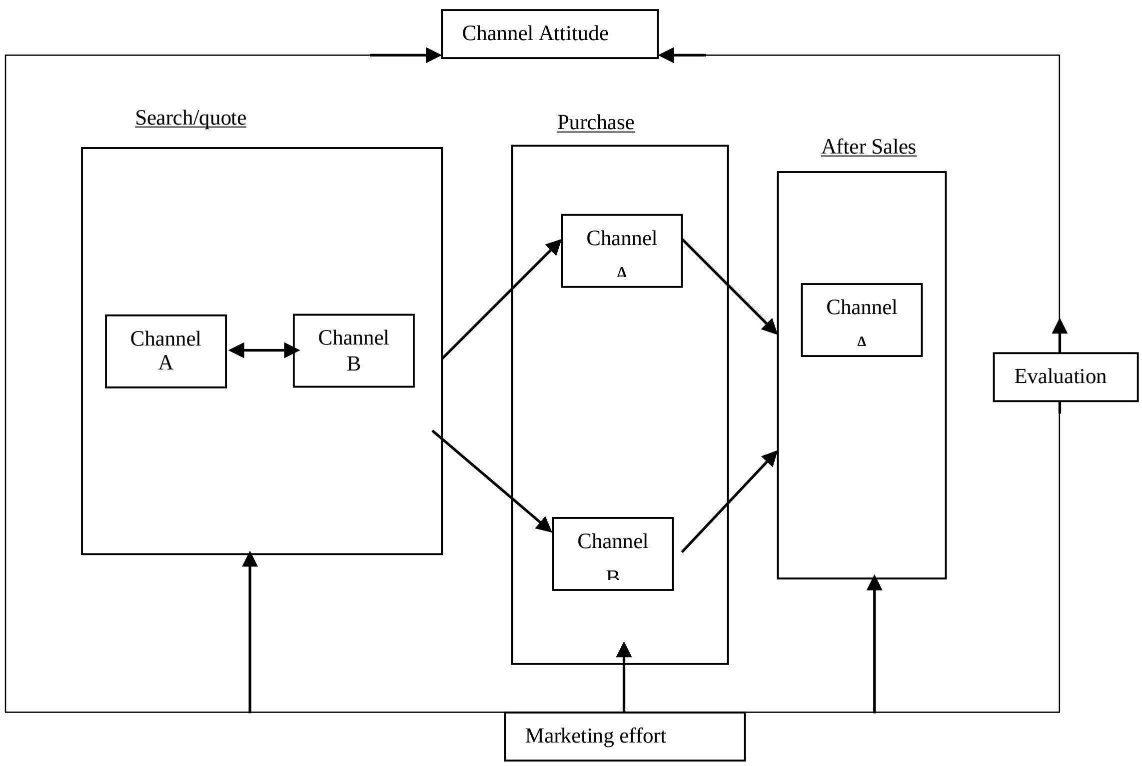
(Giaglis *et al*., 2002, pp.240)

Price comparison sites belong to the cybermediation scenario, which enables the customer to compare the prices of different car insurance companies. Coarse (1937, pp.397) found that changes in distribution 'which tend to reduce the cost of organising spatially will tend to increase the size of the firm.' This is part of the transaction cost theory which 're-intermediation and cybermediation is based on' (Jandos, 2001, pp.504). Transaction cost can cover the costs of searching for the right product, which could be lower as the customer no longer has to search through so many suppliers, to the actual purchase of the product, which is negotiable with the customer.

1.4.3. Distribution channels

Prior to aggregators, customers contacted direct insurers either by telephone or the internet. Figure1.1 represents the general model of customer choice based on the Blatterberg *et al*. (2008) general model.

Figure 1: Framework of pre-aggregator acquisition channel strategy



Key: Channel A = telephone, Channel B=own web site

The quote to sales journey from figure is as follows. Pre-aggregators, customers could either get their car insurance quotes from brokers (indirect) or from the insurance company's themselves (direct), figure 1.1 represents the direct route. Customers that use ‘direct’ can either phone the insurance company, or use their web site to gather a quote. This would entail a lengthy process to gather one quote from just one company. If the customer wanted to compare that quote, they would have to get in contact with another insurance company and repeat the process of answering many questions. Only when the customer is willing to pay for their insurance, do they proceed to the next stage of purchase. As show in figure 1.1, this can be conducted on-line or by phone. It does not matter which channel the customer originally go their quote, they could use either channel to purchase. For company x, after sales is conducted by telephone only, so for a customer to make any amendments to their policy or if they need to report a claim, this is conducted by phone, as shown in figure 1.1, after sales.

To drive people to make contact with the insurance initially, we would need to consider the marketing. Without marketing, the customer would not have contacted the company. Marketing makes customers aware of the brand and the function of the brand. Marketing affects the search/quote, purchase and after sales segments of figure 1.1. To fully understand the marketing impact, customers are asked ‘where did you hear about the company?’ This leads to the ‘**evaluation**’, which provides insight for the marketing department so that they can focus their marketing on activities that would make the biggest impact.

The final part of the diagram details channel attitude. This not only encompasses the marketing activity, but also the distribution channel of choice. This will allow company to budget their staffing levels more accurately, that there are enough staff to maintain the website, as well as to answer the important telephone calls.

1.4.4. Background to the marketing mix

The marketing mix covers many different aspects detailing the way customers perceive the company and how they conduct business with the company. People react differently to marketing, from how they perceive the company on television to their renewal reminders. The next part provides back ground to the different aspects of the marketing mix and how they influence the customer journey with respect to aggregators.

1.4.5. Marketing mix model

Price comparison sites have introduced shock effects on the business strategies and performance indicators of car insurance companies. These shocks extend into the marketing mix of the company e.g. television, radio and newspaper advertising. To measure such shocks to the marketing mix, marketing response models can be applied.

Marketing response models are not a new phenomenon with Albers reviewing, in 2012, the current state of the models. Marketing response models have not been created just for acquisition but have also been expanded to include retention (Yoo and Hanssens, 2005). Pauwels (2004) showed that long term company actions have an effect on consumers, using VAR models.

When a company experiences a shock, it has been shown that marketing can play 'an important role in turning around declining performance' of the company (Pauwels and Hanssens, 2007, pp.307). If the company were to keep its advertising the same and remain non-adaptive to its new environment, then this could make company's market share to decline and/or reduce its profitability. Marketing departments therefore need to evaluate the impact of aggregators on market performance and adapt their marketing strategies accordingly; for example, a company may choose to join the aggregators or fight them. Whichever scenario the company chooses, they must change their marketing strategy.

1.4.6. Customer retention

Customer retention plays an important role in marketing. Adding a new channel to the distribution mix can affect customer retention (Avery *et al*., 2012). Aggregators make their money every time someone purchases their car insurance via their website, so they would want customers to leave their current car insurers, at a cost to the car insurance company. This places an extra emphasis on the car insurance company to retain their customers.

As in most industries, business performance is built around generating the most value from their customers. It is generally known that retaining customers is more cost effective than acquiring new customers. The process of targeting the most profitable customer is known as customer relationship management (CRM).

The main customer segmentation targeting tool for CRM within the insurance industry revolves around customer lifetime value (CLV). Customer lifetime value within the insurance industry will not only need to consider money gained from premiums, but also claim costs. CLV is a predictive monetary value based upon customer value, stated in equation (1.1) below

$$\begin{matrix} customer \\ value \end{matrix} = \begin{matrix} Total \\ Premium \end{matrix} + \begin{matrix} Marketing \\ Costs \end{matrix} - \begin{matrix} claim \\ costs \end{matrix} \qquad (1.1)$$

To evolve this into CLV can be defined as the ‘present value of the future cash flows attributed to the customer relationship’ (Pfeifer *et al*., 2005, pp.10).

1.4.7. Introducing aggregators to the distribution channels

Reintermediation has been part of the airline industry for numerous of years. An example of how reintermedation can help a business would be to consider the official airline guide (OAG) who published the authoritative guide on airline timetables monthly. Due to the advent of the internet, they were being disintermediated from the supply chain by enabling customers to book their own flights, thus cutting out the middle man. ‘To survive, OAG had to pursue reintermedation into the supply chain of information’ (Combe, 2012, p.94). This example demonstrates the impact of companies adapting to their surroundings, to make them profitable again and grow.

The internet has bought reintermediation into many businesses worldwide. If a business operates on-line then it should consider the effects of reintermediation, but it is the least understood process in the business model (Kauffman, 2013). This does not mean that companies should not be prepared and they can either develop the essential technologies itself or procure them from present suppliers.

Aligning the business to work with a cybermediary needs to be considered with a possible change in strategy. Hiekkanen *et al.* (2013) discovered that a simple, static and mechanical approach to aggregator integration to be inadequate, and that a more dynamic approach needs to be considered. As previously mentioned, aggregators are still relatively new and so, may not behave in the traditional sense of brick and mortar businesses. This implies that businesses need to adapt to the new strategies by reviewing their current strategy to align itself with aggregators.

Online purchasing habits have been discussed before: Samaniego *et al.*, (2006), and Johnson *et al*. (1993) have discussed insurance purchasing, with Baye *et al*. (2001) discussing purchasing with the use of price comparison sites. However, neither of these studies consider the purchasing habits of car insurance customers. Thomas (2012) reviews the UK car insurance purchasing habit, but neglects the aggregator influence. Dumm and Hoyt (2003) look at the distribution channels of insurance, but their research is mainly centred on customers buying insurance on-line directly. However, Holland and Mandry

(2013) discussed the effect of price comparison sites on purchasing behaviour in numerous markets, including car insurance. They discovered that customers use price comparison sites for car insurance due to its low switching costs and low perceived consumer risk.

Appendix 4.1: The data system used in the VAR analysis

This is the rescaled data which were used in the VAR.

month	wbrat	aggrat	womrat	retrat	othrat	logspend	seasonal flag (exog3)	premind
Jan-04	-2.33	-1.69	-3.58	1.27	-3.00	13.95	1	100
Feb-04	-2.16	-1.73	-2.32	1.13	-2.80	13.96	1	99.288
Mar-04	-2.16	-1.68	-3.33	1.09	-2.63	14.08	1	98.936
Apr-04	-2.30	-1.66	-3.60	1.16	-2.53	13.88	2	103.925
May-04	-3.34	-1.49	-3.54	1.15	-2.37	14.04	2	98.865
Jun-04	-3.32	-1.36	-3.75	1.16	-2.57	13.95	2	100.65
Jul-04	-3.23	-1.59	-3.40	1.25	-2.63	13.89	3	99.954
Aug-04	-3.44	-2.24	-3.41	1.14	-2.55	14.12	3	101.122
Sep-04	-4.47	-1.78	-3.23	1.02	-2.49	14.17	3	99.965
Oct-04	-3.40	-1.46	-2.40	1.02	-2.52	13.97	4	101.151
Nov-04	-3.39	-1.33	-2.91	1.03	-2.65	13.65	4	99.822
Dec-04	-3.05	-1.21	-2.88	1.15	-2.49	13.13	4	101.564
Jan-05	-2.48	-2.09	-1.70	1.10	-2.45	13.22	1	102.377
Feb-05	-2.52	-3.83	-1.77	1.02	-2.49	13.33	1	101.967
Mar-05	-2.55	-4.04	-1.77	0.94	-2.49	13.38	1	102.021
Apr-05	-2.51	-4.10	-1.79	0.87	-2.53	13.13	2	103.475
May-05	-2.63	-4.44	-1.85	0.87	-2.57	13.11	2	100.571
Jun-05	-2.63	-4.53	-1.81	0.85	-2.62	13.04	2	100.639
Jul-05	-2.73	-4.28	-1.82	0.81	-2.63	13.21	3	99.166
Aug-05	-2.59	-4.20	-1.72	0.80	-2.57	13.36	3	96.861
Sep-05	-2.49	-3.97	-1.63	0.82	-2.54	13.39	3	99.033
Oct-05	-2.47	-3.72	-1.64	0.89	-2.50	13.14	4	100.615

Nov-05	-2.43	-3.99	-1.62	0.87	-2.47	12.92	4	101.281
Dec-05	-2.55	-4.17	-1.67	1.00	-2.53	12.75	4	100.488
Jan-06	-2.63	-3.20	-1.80	0.99	-2.61	13.19	1	99.34
Feb-06	-2.76	-2.60	-1.75	0.84	-2.60	13.30	1	100.04
Mar-06	-2.75	-2.65	-1.72	0.81	-2.56	13.36	1	100.603
Apr-06	-2.47	-2.38	-1.61	0.87	-2.50	13.26	2	98.945
May-06	-2.66	-2.38	-1.58	0.92	-2.49	13.11	2	97.926
Jun-06	-2.58	-2.46	-1.50	0.92	-2.40	13.14	2	100.893
Jul-06	-2.55	-2.54	-1.48	1.01	-2.44	13.16	3	97.955
Aug-06	-2.64	-2.39	-1.56	0.96	-2.46	13.08	3	99.753
Sep-06	-2.54	-2.36	-1.46	0.91	-2.39	13.25	3	101.439
Oct-06	-2.56	-2.38	-1.47	0.91	-2.36	13.33	4	100.152
Nov-06	-2.44	-2.55	-1.54	0.95	-2.43	12.87	4	103.887
Dec-06	-2.53	-2.52	-1.42	1.02	-2.37	12.40	4	100.486
Jan-07	-2.39	-2.55	-1.55	1.10	-2.41	13.44	1	95.967
Feb-07	-2.44	-2.51	-1.62	0.92	-2.43	13.66	1	100.121
Mar-07	-2.39	-2.52	-1.51	0.91	-2.37	13.79	1	102.309
Apr-07	-2.26	-2.43	-1.43	0.94	-2.31	13.16	2	98.989
May-07	-2.43	-3.99	-1.46	0.93	-2.36	13.18	2	96.663
Jun-07	-2.60	-4.36	-1.48	0.94	-2.40	12.94	2	102.532
Jul-07	-2.72	-4.41	-1.50	0.85	-2.41	12.60	3	98.67
Aug-07	-2.89	-4.32	-1.51	0.85	-2.47	13.06	3	93.607
Sep-07	-2.78	-3.67	-1.49	0.84	-2.45	13.20	3	99.064
Oct-07	-2.78	-3.77	-1.48	0.87	-2.47	12.66	4	103.841
Nov-07	-2.65	-3.75	-1.50	0.88	-2.47	11.54	4	100.126

Dec-07	-2.40	-3.84	-1.33	0.96	-2.49	11.76	4	104.383
Jan-08	-2.75	-3.88	-1.59	0.99	-2.58	12.59	1	90.409
Feb-08	-2.80	-3.85	-1.36	0.83	-2.42	12.67	1	102.967
Mar-08	-2.90	-4.08	-1.79	0.84	-2.68	12.59	1	100.716
Apr-08	-3.15	-4.05	-1.68	0.82	-2.58	12.56	2	103.372
May-08	-3.18	-3.96	-1.60	0.88	-2.57	12.49	2	97.192
Jun-08	-2.88	-3.81	-1.66	0.90	-2.55	12.05	2	97.519
Jul-08	-3.04	-3.73	-1.70	0.95	-2.53	12.60	3	97.79
Aug-08	-2.84	-3.81	-1.78	0.96	-2.64	12.82	3	102.775
Sep-08	-2.98	-4.10	-1.76	0.86	-2.67	13.15	3	106.137
Oct-08	-2.80	-4.31	-1.85	0.76	-2.67	12.52	4	101.527
Nov-08	-2.78	-4.41	-1.83	0.73	-2.71	11.64	4	104.334
Dec-08	-2.76	-4.53	-1.84	0.80	-2.68	11.24	4	100.576
Jan-09	-2.80	-4.74	-2.12	0.86	-2.85	12.41	1	89.105
Feb-09	-3.05	-4.78	-2.16	0.75	-2.99	12.41	1	101.293
Mar-09	-3.20	-4.72	-2.13	0.78	-2.91	11.77	1	103.06
Apr-09	-3.18	-4.82	-2.12	0.78	-2.92	11.35	2	103.086
May-09	-2.95	-4.92	-2.16	0.78	-3.01	10.88	2	102.115
Jun-09	-3.28	-4.88	-2.14	0.71	-2.98	11.75	2	100.099
Jul-09	-3.28	-4.90	-2.14	0.77	-3.06	11.55	3	98.752
Aug-09	-3.47	-4.79	-2.17	0.73	-3.05	11.79	3	96.073
Sep-09	-3.25	-5.50	-2.19	0.78	-3.10	10.20	3	109.38
Oct-09	-3.44	-6.04	-2.39	0.82	-3.25	11.53	4	106.024
Nov-09	-3.36	-6.15	-2.39	0.80	-3.24	10.07	4	100.424
Dec-09	-3.32	-6.21	-2.22	0.95	-3.37	9.17	4	103.956

Appendix 4.2: Complete VAR model

The results can be written as a 4th order Vector autoregressive model

$$yt = \begin{pmatrix} 0.491 & 0.519 & 0.509 & 0.378 & 0.525 & 1.094 \\ 0.041 & 0.803 & 0.407 & 0.233 & 0.511 & 1.059 \\ 0.400 & 0.450 & 0.754 & 0.579 & 0.560 & 0.996 \\ 0.851 & 0.361 & 0.536 & 0.675 & 0.481 & 0.800 \\ 0.368 & 0.356 & 0.170 & 0.475 & 0.930 & 1.388 \\ 95.337 & 1.517 & 1.036 & 0.094 & 1.325 & 4.052 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0.635 & 0.482 & 0.553 & 0.465 & 0.475 & 1.057 \\ 0.350 & 0.607 & 0.554 & 0.626 & 0.488 & 1.207 \\ 0.330 & 0.419 & 0.716 & 0.194 & 0.447 & 1.213 \\ 0.534 & 0.480 & 0.461 & 0.416 & 0.511 & 1.071 \\ 0.498 & 0.731 & 0.402 & 0.049 & 0.408 & 0.436 \\ 23.420 & 1.166 & 0.415 & 0.326 & 1.435 & 0.662 \end{pmatrix} y_{t-2}$$

$$+ \begin{pmatrix} 0.558 & 0.488 & 0.462 & 0.634 & 0.523 & 1.099 \\ 0.272 & 0.693 & 0.602 & 0.622 & 0.562 & 1.132 \\ 0.211 & 0.716 & 0.466 & 0.957 & 0.467 & 1.379 \\ 0.492 & 0.559 & 0.386 & 0.494 & 0.523 & 0.921 \\ 0.879 & 0.475 & 0.501 & 0.997 & 0.527 & 1.008 \\ 40.995 & 0.581 & 1.422 & 83.383 & 0.678 & 0.373 \end{pmatrix} y_{t-3} + \begin{pmatrix} 0.310 & 0.500 & 0.486 & 0.365 & 0.501 & 1.021 \\ 0.584 & 0.420 & 0.277 & 0.148 & 0.476 & 1.208 \\ 0.448 & 0.556 & 0.205 & 0.180 & 0.500 & 1.131 \\ 0.472 & 0.516 & 0.555 & 0.335 & 0.503 & 0.882 \\ 0.198 & 0.502 & 0.738 & 0.079 & 0.471 & 1.917 \\ 0.615 & 0.459 & 2.298 & 0.027 & 0.700 & 1.891 \end{pmatrix} y_{t-4}$$

Appendix 6.1: Quantile value model model

Parameter		DF	Estimate
Intercept		1	189.7399
claim	Y	1	-375.105
claim	|N	0	0
paymthd	D	1	129.0165
paymthd	S	1	-3.1734
paymthd	|C	0	0
ncbp	P	1	-29.5412
ncbp	|N	0	0
cover	F	1	-76.5819
cover	T	1	-82.9794
cover	|C	0	0
region	EAST ANGLIA	1	-62.4653
region	EAST MIDLANDS	1	-4.9387
region	GREATER LONDON	1	122.8923
region	N Ireland	1	193.6547
region	NORTH	1	27.3328
region	NORTH WEST	1	151.2435
region	OTHER	1	-183.706
region	SCOTLAND	1	-116.95
region	SOUTH WEST	1	-42.731
region	WALES	1	-4.0252
region	WEST MIDLANDS	1	70.0565
region	YORKSHIRE & HUMBER	1	68.3316
region	_OTHER SOUTHEAST	0	0
sex	F	1	-35.9245
sex	M	0	0
married	C	1	-53.7916
married	M	1	-43.316
married	O	1	30.102
married	S	0	0
agegrp	23 and u	1	474.3673
agegrp	24-28	1	206.8564
agegrp	29-32	1	73.4504
agegrp	33-37	1	29.4039
agegrp	46 +	1	-37.5446

agegrp	|38-46	0	0
ncbgrp	1	1	-26.8315
ncbgrp	11	1	-253.892
ncbgrp	4 & 5	1	-140.914
ncbgrp	6 to 8	1	-161.086
ncbgrp	9 to 10	1	-174.617
ncbgrp	|2 - 3	0	0
rnyear	1	1	344.2555
rnyear	2	1	729.2206
rnyear	3+	1	1597.208
rnyear	|0	0	0
vehgrpa	01-Apr	1	-118.852
vehgrpa	13-20	1	167.3399
vehgrpa	21+	1	65.4992
vehgrpa	05-Jun	1	-71.7648
vehgrpa	7 to 12	0	0
caragea	01-Mar	1	22.7443
caragea	10+	1	-97.9235
caragea	3 to 5	1	33.5359
caragea	|6 to 9	0	0
numdrv	1	1	283.387
numdrv	2	1	217.9074
numdrv	3	1	261.5445
numdrv	4	1	327.915
numdrv	5	0	0
media	DIRECTORIES	1	116.3737
media	DOOR DROP	1	6.7505
media	ONLINE	1	-30.1211
media	OUTDOR	1	-6.4944
media	PRESS & MAGS	1	160.2784
media	RADIO	1	743.0759
media	REFERRAL	1	37.9685
media	TV	1	-16.2212
media	UNKNOWN	1	46.2141
media	WIN BACK	1	-58.8902
media	|AGGREGATOR	0	0

Appendix 6.2: General linear model for value

Parameter	Estimate
Intercept	380.6478
claim Y	-1988.64
claim |N	0
paymthd D	210.0517
paymthd S	19.05302
paymthd |C	0
ncbp P	-98.6255
ncbp |N	0
cover F	-141.75
cover T	-253.891
cover |C	0
region EAST ANGLIA	-37.1644
region EAST MIDLANDS	15.24081
region GREATER LONDON	176.8259
region N Ireland	391.9792
region NORTH	-40.7887
region NORTH WEST	65.11802
region OTHER	-356.382
region SCOTLAND	-176.178
region SOUTH WEST	-161.856
region WALES	-48.6032
region WEST MIDLANDS	-30.2072
region YORKSHIRE & HUMBER	41.20613
region _OTHER SOUTHEAST	0
sex F	-112.452
sex M	0
married C	-121.595
married M	-68.1263
married O	78.6697
married S	0
agegrp 23 and u	604.7882
agegrp 24-28	288.661
agegrp 29-32	127.2373
agegrp 33-37	-14.8512
agegrp 46 +	-144.025
agegrp |38-46	0
ncbgrp 1	-528.752

ncbgrp 11	-177.041
ncbgrp 4 & 5	-91.5309
ncbgrp 6 to 8	-71.7718
ncbgrp 9 to 10	-149.375
ncbgrp |2 - 3	0
rnyear 1	308.8998
rnyear 2	792.659
rnyear 3+	1741.701
rnyear |0	0
vehgrpa 1-4	-191.673
vehgrpa 13-20	237.1669
vehgrpa 21+	108.9171
vehgrpa 5-6	-107.544
vehgrpa 7 to 12	0
caragea 1-3	-27.3729
caragea 10+	-196.697
caragea 3 to 5	11.18721
caragea |6 to 9	0
numdrv 1	351.0496
numdrv 2	236.9629
numdrv 3	348.1342
numdrv 4	462.9861
numdrv 5	0
media DIRECTORIES	137.9676
media DOOR DROP	-87.4109
media ONLINE	-115.988
media OUTDOR	-59.5855
media PRESS & MAGS	100.6708
media RADIO	-169.114
media REFERRAL	54.01172
media TV	-92.2578
media UNKNOWN	-127.845
media WIN BACK	-163.521
media |AGGREGATOR	0

Appendix 6.3: Winzorised general linear model for value

Parameter	Estimate
Intercept	75.47076
claim Y	-948.861
claim |N	0
paymthd D	161.3256
paymthd S	-14.7879
paymthd |C	0
ncbp P	-73.6878
ncbp |N	0
cover F	-125.26
cover T	-194.22
cover |C	0
region EAST ANGLIA	-69.2139
region EAST MIDLANDS	-9.29983
region GREATER LONDON	181.0964
region N Ireland	370.1469
region NORTH	24.83303
region NORTH WEST	214.813
region OTHER	-328.633
region SCOTLAND	-189.348
region SOUTH WEST	-79.865
region WALES	-17.9643
region WEST MIDLANDS	79.97422
region YORKSHIRE & HUMBER	107.9659
region _OTHER SOUTHEAST	0
sex F	-85.229
sex M	0
married C	-91.8346
married M	-61.7587
married O	101.2182
married S	0
agegrp 23 and u	601.9858
agegrp 24-28	286.7092
agegrp 29-32	105.7385
agegrp 33-37	41.80326
agegrp 46 +	-56.0804

agegrp |38-46	0
ncbgrp 1	-191.717
ncbgrp 11	-231.124
ncbgrp 4 & 5	-123.518
ncbgrp 6 to 8	-87.3424
ncbgrp 9 to 10	-154.186
ncbgrp |2 - 3	0
rnyear 1	380.08
rnyear 2	831.9324
rnyear 3+	1721.531
rnyear |0	0
vehgrpa 1-4	-158.204
vehgrpa 13-20	266.0391
vehgrpa 21+	126.4792
vehgrpa 5-6	-93.0201
vehgrpa 7 to 12	0
caragea 1-3	19.28675
caragea 10+	-139.536
caragea 3 to 5	39.11153
caragea |6 to 9	0
numdrv 1	466.6502
numdrv 2	383.3503
numdrv 3	484.3117
numdrv 4	555.6798
numdrv 5	0
media DIRECTORIES	99.39804
media DOOR DROP	-170.071
media ONLINE	-94.1058
media OUTDOR	-43.7901
media PRESS & MAGS	136.2666
media RADIO	499.9187
media REFERRAL	24.43709
media TV	-51.172
media UNKNOWN	22.77149
media WIN BACK	-124.827
media |AGGREGATOR	0

Appendix 6.4: Logistic Renewal Model

Parameter		DF	Estimate
Intercept		1	2.2348
Age group	23 and under	1	0.0619
Age group	24-28	1	0.0367
Age group	29-32	1	0.078
Age group	33-37	1	0.0216
Age group	46 +	1	-0.1445
Car age	1 to 3	1	-0.1877
Car age	10+	1	0.1748
Car age	3 to 5	1	-0.0725
Claimed on insurance	Y	1	-0.1543
Social grouping	A) Wealthy Executives	1	0.00183
Social grouping	B) Affluent Greys	1	-0.1791
Social grouping	C) Flourishing Families	1	-0.0808
Social grouping	D) Prosperous Professionals	1	0.2414
Social grouping	E) Educated Urbanites	1	0.3992
Social grouping	F) Aspiring Singles	1	0.1762
Social grouping	G) Starting Out	1	0.16
Social grouping	I) Settled Surburbia	1	-0.0755
Social grouping	J) Prudent Pensioners	1	0.07
Social grouping	K) Asian Communities	1	0.161
Social grouping	L) Post-Industrial Families	1	-0.1302
Social grouping	M) Blue Collar Roots	1	-0.0812
Social grouping	N) Struggling Families	1	-0.2008
Social grouping	O) Burdened Singles	1	-0.3632
Social grouping	P) High-Rise Hardship	1	-0.008
Social grouping	Q) Inner City Adversity	1	-0.0973
Social grouping	R) Communal & Others	1	0.1216
Social grouping	Unknown	1	-0.00057
Marketing Source	DIRECTORIES	1	-0.0165

Marketing Source	DOOR DROP	1	-1.1339
Marketing Source	ONLINE	1	0.0127
Marketing Source	OUTDOR	1	0.3855
Marketing Source	PRESS & MAGS	1	0.1572
Marketing Source	RADIO	1	0.4703
Marketing Source	REFERRAL	1	-0.0662
Marketing Source	TV	1	0.0786
Marketing Source	UNKNOWN	1	-0.0237
Marketing Source	WIN BACK	1	0.037
No Claims Bonus	1	1	-0.0483
No Claims Bonus	11	1	0.1223
No Claims Bonus	4 & 5	1	-0.2024
No Claims Bonus	6 to 8	1	-0.0618
No Claims Bonus	9 to 10	1	0.1387
No Claims Bonus protected	P	1	-0.044
Number of drivers	1	1	-1.3753
Number of drivers	2	1	-1.576
Number of drivers	3	1	-1.4316
Number of drivers	4	1	-1.4511
Allowed to contact policy holder	N	1	-0.1263
Pay method	D	1	0.1105
Pay method	S	1	-0.0559
UK region	EAST ANGLIA	1	-0.1548
UK region	EAST MIDLANDS	1	0.0442
UK region	GREATER LONDON	1	0.1372
UK region	N Ireland	1	-0.3259
UK region	NORTH	1	-0.2444
UK region	NORTH WEST	1	-0.0119
UK region	OTHER	1	0.3926
UK region	SCOTLAND	1	0.2761
UK region	SOUTH WEST	1	-0.00062
UK region	WALES	1	-0.2005
UK region	WEST MIDLANDS	1	-0.1245
UK region	YORKSHIRE & HUMBER	1	0.093
Renewal year	1	1	-0.1775

Renewal year	2	1	0.0447
Renewal year	3+	1	0.3226
Gender	F	1	-0.1113
Value		1	0.000015
Vehicle group	01-04	1	-0.0641
Vehicle group	13-20	1	0.022
Vehicle group	21+	1	0.022
Vehicle group	05-06	1	-0.00859

Appendix 7.1: Data for graphs

Month	ACR	WOM sales	Ambient Spend	Direct Sales	web sales	web spend	DM quote	DM spend	Retained
May-07	1.81%	5273	2.81	3.22	2.60	1.92	2.09	1.92	72%
Jun-07	1.26%	4768	2.66	2.53	2.25	1.15	2.36	2.06	72%
Jul-07	1.20%	4924	0.23	2.47	2.25	2.13	2.71	1.87	70%
Aug-07	1.31%	4141	2.59	1.89	1.76	1.65	2.53	2.32	70%
Sep-07	2.47%	3637	4.18	1.64	1.52	0.94	1.58	1.88	70%
Oct-07	2.25%	3684	1.97	1.43	1.42	0.91	1.63	2.86	71%
Nov-07	2.31%	3522	0.11	1.31	1.12	1.11	1.45	1.41	71%
Dec-07	2.10%	2828	0.10	0.90	0.89	0.92	0.76	1.11	72%
Jan-08	2.02%	4093	0.90	1.38	1.42	1.52	1.43	1.11	73%
Feb-08	2.09%	3630	1.14	1.19	1.17	1.50	0.87	1.15	70%
Mar-08	1.66%	3587	1.18	1.18	1.22	1.27	1.09	0.90	70%
Apr-08	1.71%	3720	1.24	1.21	1.28	1.11	1.13	1.57	69%
May-08	1.86%	3858	0.32	1.11	1.33	1.75	0.99	1.20	71%
Jun-08	2.16%	3562	0.36	1.08	1.23	1.78	0.94	1.05	71%
Jul-08	2.34%	3648	0.61	1.12	1.21	1.79	1.03	0.82	72%
Aug-08	2.05%	2950	1.63	0.90	1.02	1.58	0.81	1.38	72%
Sep-08	1.59%	1570	2.18	0.51	0.56	1.57	0.48	0.68	70%
Oct-08	1.33%	1237	0.53	0.41	0.50	0.68	0.52	0.40	68%
Nov-08	1.20%	1108	0.42	0.31	0.40	0.53	0.38	0.26	67%
Dec-08	1.06%	906	0.43	0.25	0.33	0.37	0.24	0.25	69%
Jan-09	0.86%	969	0.55	0.31	0.41	0.55	0.39	0.30	70%
Feb-09	0.83%	823	0.72	0.25	0.31	0.42	0.39	0.18	68%
Mar-09	0.88%	934	0.10	0.29	0.36	0.44	0.42	0.25	69%
Apr-09	0.80%	777	0.30	0.24	0.29	0.10	0.43	0.30	69%
May-09	0.72%	683	0.17	0.20	0.29	0.08	0.35	0.20	69%
Jun-09	0.76%	699	0.54	0.24	0.29	0.08	0.30	0.18	67%
Jul-09	0.74%	703	0.42	0.22	0.30	0.08	0.39	0.18	68%
Aug-09	0.82%	642	0.55	0.21	0.27	0.09	0.33	0.21	68%