Customer Retention and Customer Complaints: An Empirical Analysis of Two Subscription-based Products

Author:
Naina Mohamed, Jeenat Beham

Publication Date:
2015

DOI:
https://doi.org/10.26190/unsworks/2936

Customer Retention and Customer Complaints: An Empirical Analysis of Two Subscription-based Products

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Marketing

Jeenat Beham Naina Mohamed

The School of Marketing
UNSW Business School
The University of New South Wales
Sydney, Australia
September, 2015

ABSTRACT

This research is undertaken within the field of relationship marketing. The study focuses on customer retention with emphasis on understanding customer defection in the context of customer complaint behaviour. The study examines two telecommunication products, Internet services and voice services. These represent two telecommunications products, which are at the growth and decline stages of the product life cycle, respectively.

A number of explanatory variables are examined in detail with the objective of exploring the effect on customer retention, in terms of relationship strength and relationship direction. This study extends existing studies of retention by not only looking at explanatory variables that have been studied in previous research, but extending the range of explanatory variables to focus on a number of key aspects related to customer complaint behaviour. These additional variables include a) whether a customer ever complained, b) the number of complaints, c) the time since a customer last made a complaint, d) severity of the complaints, and e) length of the recovery from the complaint.

The study involves an analysis of longitudinal data, using survival analysis, for 15,000 to 20,000 customers over three years. Three increasingly descriptive models are estimated for each product: a baseline model that utilises customer characteristics to explain retention; a comprehensive model including time-varying explanatory variables and finally, possible interactive effects based on the comprehensive model are evaluated. Factors such as non-complainants, length of recovery, communication encounters, household size and household income are found to contribute to an *increased risk of defection*. On the other hand, the number of complaints, high severity incidents, 'recency' of complaints, usage and, length of the relationship are found to contribute to a *decreased* *risk of defection*. Compared with previous studies, additional depth of analyses is provided through a comparison of products at different stages of the product life-cycle and the varying descriptors presented.

This research has important theoretical and practical implications. The study identifies that using a combination of customer characteristics, customer transaction behaviour and customer complaint provides a much greater understanding of customer retention behaviour.

Keywords:

Customer complaining behaviour, customer retention; customer defection, customer churn, service failure, service recovery, customer characteristic, survival analysis, relationship marketing.

ACKNOWLEDGEMENTS

The completion of this dissertation causes me to reflect on my tenure as a PhD candidate. This longest of journeys would have not been possible without God's grace and the profound influence of loved ones, inspirational friends and colleagues.

For my supervisors, Professor Adrian Payne and Dr. Rahul Govind, I have gratitude and praise. It has been a privilege to work with them. Their advice and feedback has been invaluable to me in developing my research skills. They helped me strive to submit the very best dissertation I could write. Special thanks also go to the many expert faculty members for their stimulating advice, suggestions and encouragement: Professor John Roberts, Professor Ashish Sinha, Professor Paul Patterson, Professor Mark Uncles, and Associate Professor Liem Viet Ngo. Thank you one and all for being unstinting with your time. A special thank you also goes to Associate Professor Dr. Jack Cadeaux and the faculty of the School of Marketing for enabling me to fulfil my dream. Special thanks also to our administration staff Nadia Withers, Paula Aldwell and Margot DeCelis for all of their help throughout my studies.

Alongside the acknowledgements to the university, I owe a debt of gratitude to my family and my Sydney family for their constant encouragement and providing moral and spiritual support. In particular, I am forever indebted to my dear husband Thajuddin, for setting aside his career for mine and then standing by me through every step of this tempestuous journey. I could have not completed this dissertation without his enduring love, sacrifice and compassion. Special thanks also to my daughters Irdina and Isha for their treasured insights and unconditional loving support.

I also acknowledge the support of my fellow students and friends who have encouraged me, helped me survive the stress and develop my sticking power, pushing me to achieve more than I thought I could. My close friend, Kaye Chan for her academic and personal support; for sharing and being part of this journey. My dear friend, Dr. Marion Bufford, for her guiding wisdom and practical insights. I am also thankful for the goodwill of so many other good friends - Room 3005 friends (David Lie, Ngoc Lu, Yu Tian, Christopher Agyapong Siaw), Room 3003 friends (Eileen Chiew, Rawi Roongruangsee, Jake An, Rebecca Scott), and my lovely Nicole Lasky. Thank you all for helping me out and being there for me.

Last but not the least; I am indebted to my sponsors, *YTM*, for providing me with the necessary resources and financial support to complete this dissertation.

I say sincerely thank you one and all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.0 Introduction

The aim of this research is to examine the factors impacting customer retention by focusing on complaint behaviour. Limited studies have focused on the factors influencing customer retention, especially with respect to complaint behaviour. Further, extant customer retention research tends to use perceptual metrics and provides only a narrow understanding of customer retention behaviour over time.

Studies demonstrate that customer complaint behaviour is a crucial dimension of the customer-firm relationship and customer retention (Fornell & Wernerfelt, 1988). However, scant attention has been paid to understanding retention using observable metrics (Braun & Schweidel, 2011). Limited studies (Knox & Van Oest, 2014; Lariviere & Van den Poel, 2005) have utilised customer complaint behaviour metrics to help understanding retention.

This study, in addition to considering previously identified drivers of customer retention such as customer characteristics, use, and length of relationship, uses longitudinal data over a period of three years to capture the simultaneous effects of several explanatory variables using a dynamic statistical approach. Traditional drivers of customer retention models using customer satisfaction, retention, and transaction history provide somewhat limited results. This study extends existing work, testing the robustness of previous modelling and identifying the usefulness of some new explanatory variables. The study focuses on customers subscribing to telecommunications products in the decline and growth phases of the product life cycle. This thesis extends previous complaint studies using real complaint data in a longitudinal model.

1.1 Research Focus and Domain of Study

This research uses relationship marketing (RM) as the broad domain of academic knowledge. Within this domain it focuses on making a contribution to the customer retention and customer complaint literature. The shift from transactional to relationship marketing represented a critical shift in marketing theory as academics and practitioners recognised, especially in services industries, that most exchanges were relational (Berry, 1983; Dwyer, Schurr, & Oh, 1987; Grönroos, 1994; Gummesson, 1994). Since this shift, academic RM research has increased substantially. One area of particular importance is customer retention.

1.2 Customer Retention and Defection

This study focuses on the retention within the rapidly growing telecommunications sector. Telecommunications firms, which provide multiple subscription-based services, including Internet, voice, and cable TV services, are facing significant challenges in profitability, competition, buyer power, and barriers to entry. All of these factors affect margins and market share. An important factor is the need to focus on the customer retention to sustain recurring revenue (Braun & Schweidel, 2011; Petersen et al., 2009).

Existing studies focus on linking retention rates to customer lifetime value (Fader & Hardie, 2010), customer equity management (Rust, Lemon, & Zeithaml, 2004), resource allocation (Reinartz, Thomas, & Kumar, 2005; Reinartz & Kumar, 2003), and financial reporting and management (Gupta, Lehmann, & Stuart, 2004). Gupta and Zeithaml (2006) define customer retention as “the probability of a customer being ‘alive,’ or repeat buying from a firm.” Ahmad and Buttle (2001) define customer retention as “the mirror image of customer churn, where a high retention rate has the same significance as a low churn rate. In this study, customer churn and defection are used interchangeably to refer to a lost customer (Ahmad & Buttle, 2001).

Much research focuses on understanding processes leading to customer retention but provide limited insights into building retention using customer complaint behaviour. While most firms monitor complaints, the key factors influencing customer retention and defection are not always well understood. Today, the markets for many services have reached maturity or decline. Declining markets require special attention to be placed on understanding not only the value of customer retention, but also the negative impact of customer defection on profitability and brand value.

1.3 Importance of Complaint Management

The telecommunications industry has been one of the fastest growing industries over the past few decades (OECD Science, Technology and Industry Outlook, 2014). While many companies within the sector have focused on the brand proposition and strengthening brand loyalty (Bolton, Lemon, & Verhoef, 2002), the essence of a brand is its reliability. It is important to recognise that customer complaints represent an important area on which to focus in order to improve reliability. Effectively managing customer complaints represent an integral component of maintaining and improving market standing.

In the last two decades in particular, deregulation has caused a surge in competition among telecommunications, utilities, and other service industries. The telecommunications industry, in particular, is highly competitive, mature, and plagued by churn (Ahn, Han, & Lee, 2006). This pattern is especially prevalent in countries where the government has permitted wireless number portability, removing the main obstacle to changing mobile carriers (e.g., Wong, 2011). Further, consumers in some metropolitan areas of more developed countries can have a choice of more than 10 mobile phone carriers. Telecommunications customers are easily prompted to defect over dropped calls and service failures (Ahn et al., 2006; Chandrashekaran, Rotte, Tax, & Grewal, 2007). As a result, how complaints are handled is critical to retaining customers and is of growing concern to businesses in this industry. This is especially true for high-volume businesses as repeated unsatisfactory service can be highly detrimental to customer retention (Alvarez, Casielles, & Martin, 2010; Chandrashekaran et al., 2007).

It is well established that dissatisfied customers may not complain often, but rather, dissatisfaction is detected through usage patterns. Service failures may or may not significantly impact customer defection behaviour as service failure recovery time plays a role in defection behaviour. Differing telecommunications service plan tiers also play a role, as customers seek differing levels of optimisation of their call usage. Therefore, understanding complaint management is critical.

1.4 Customer Expectations in Service Economies

The trend of both developed and developing countries, who are moving towards service-based economies, suggests that customers expect and demand more than just competent services from the service providers. Rather, they expect a level of customisation and differentiation, which is provided through appropriate levels of services (Reinartz & Ulaga, 2008). The employment and trade produced by service-based economies comprises two-thirds of the world's Gross Domestic Product (OECD World Economic Outlook Annual Report, 2012). Total world trade in goods and commercial services has increased 20 per cent since 2010. These services include communications, insurance, and financial services. Today, increasingly firms are moving toward a subscription-based model. This can be seen in firms such as Netflix, Amazon's Kindle, cable companies, and cell phone companies. Gartner industry[1] research firm predicts that by 2015, 35 per cent of all firms will generate revenue via subscription-based services and recurring revenue models.

1.5 Subscription-based Services

Subscription services derive sales from recurring revenue, including upgrades, add-ons, premium packages, and product bundles. Once a critical mass of existing customers is reached, firms focus on providing ongoing value to customers to ensure recurrent revenue stream models. Such a model follows consumers moving from one-time purchases to subscription services. For marketers, this shift opens the door to a new set of revenue opportunities, but it also requires fundamental changes to how they market services. It also increases the necessity of targeted marketing strategies complementing demographic shifts, as opposed to aggregative undifferentiated marketing strategies.

---

[1] The information is retrieved is sourced from Gartner Inc. subscribed by Firm A. The report is part of copyrighted Industry Report & Forecasts Series dated January 2010.

Firms must understand the mechanics of retention and defection of customers. An interest in strategies specifically aimed at lowering customer defection rates (or conversely, increasing customer retention rates) emerged in management sciences during the 1990s led by Reichheld and Sasser (1990) and Reichheld (1996). While some, such as Dowling and Uncles (1997) and Reinartz and Kumar (2000) have argued that the link between customer retention and profitability is not clear-cut, there is a general consensus that preventing customer defection is a sound business strategy (Hogan, Lemon, & Libai, 2003). Further, some argue that overlooking customer defection is among the worst sources of profit leakage (Bhote & Bhote, 2004).

Current evidence suggests that studies that determine the causes of customer defection prevent revenue loss for a firm; representing a significant opportunity for profit improvement in highly competitive environments. However, despite several decades of advancement in this field, defection rates across the world and across industries remain high. Blattberg, Kim, and Neslin (2008) and Boone and Kurtz (2011) have shown that it is common for industries to have annual defection rates between 20 and 50 per cent.

Further, it typically costs a firm more to find new customers than it does to retain current ones. Loyal customers are likely to increase the depth or breadth of their relationship with a firm over time if the firm provides high-quality service. Additionally, rising defection rates increase the average unit cost and decrease profitability per customer. Reichheld and Sasser (1990) show that even a five per cent points improvement in retention rates led to an increase in overall profit of 75 per cent in the credit card industry, and 35 per cent in the computer software industry. Other researchers have since reached similar conclusions (Gupta et al., 2004; Winer, 2001).

1.6 Different Stages of the Product Life Cycle

Different product growth characteristics and different stages of the product life cycle raise an interesting question. Is customer retention behaviour similar or different for customers subscribing to products in the decline stage versus the growth stage of the life-cycle? Research has shown that customer retention has a substantial impact on a growing market (Hogan et al., 2003; Thompson & Sinha, 2008). Yet, limited attempts have been made to include such market characteristics in extant retention studies. One major limitation is the combined lack of sufficient longitudinal data (Gupta et al., 2004) and lack of customer touch points information for these types of services (Van den Bulte & Iyengar, 2011). Including this type of information should provide further insight into customer retention behaviour. Comparing retention behaviour across these two different stages of growth and decline should not only help manage products more proactively, but also help allocate resources appropriately, depending on the stages of products' life cycles.

1.7 Motivation and Aim of Research

Global changes in how complaints are handled have led to growing concerns among firms. This warrants consideration of new models for analysing behavioural patterns of customers. Traditional models do not typically represent behavioural patterns well enough for firms to understand their positions in competitive markets. This study looks into existing theoretical elements of complaint management and considers models relevant to the motivations and aims of this research. The key motivations of this research are as follows:

First, defection rates are rising among service providers. Although there are many studies focused on understanding retention, defection rates across the services industry, according to one study, cause revenue losses of USD 4 billion per annum (Allison, 2010). For example, according to EDF Energy, high levels of defection characterise the UK energy market. At least 160,000 customers defect to an alternative electricity or gas supplier every week - an average defection rate of 38 per cent, which equates to approximately £300 million a year[2].

Second, data used in retention modelling is often incomplete. Initial retention models were built around survey data, either using random surveys of customer satisfaction or "exit surveys" (Wong, 2011). However advances in information technology and data storage in recent years have opened the way for models to be built on longitudinal data constructed from a firm's client records; as a result firms now have an opportunity to use more operationally-oriented metrics, allowing for investigation of retention trigger factors, including service quality and customer complaints (Braun & Schweidel, 2011).

Third, marketing teams must enhance their skills. Common views of retention are often short-sighted. This is because research teams have limited data to assess. Even when ample data is available, marketing teams may be unable to efficiently and effectively analyse it due to a lack of skills and knowledge and/or the inability to analyse huge amounts of data. This research aims at supporting marketing managers' needs to better understand the crucial factors impacting customer retention and defection.

[2] The information is retrieved from:
http://www.sas.com/content/dam/SAS/en_gb/doc/CustomerStories/edf.pdf

Finally, the volume of data is growing exponentially. More than 15 billion devices are expected be connected to the Internet in 2015[3]. This increase in data permits investigation of how multiple potential triggers, such as service problems or customer usage history, may jointly contribute to a customer's defection behaviour.

1.8 Key Gaps Addressed in the Literature

A number of gaps in the extant literature are addressed in this research. First, although complaint management is an important aspect of RM, most studies are limited by their reliance on surveys and perceptual measures, such as repurchase intentions and satisfaction, as surrogate measures of retention. Perceptual measures of retention may not be an accurate surrogate for actual defection behaviour (Mittal & Kamakura, 2001). Further, studies in customer retention typically deal with few specific factors, as opposed to using integrated, more comprehensive models to examine customer retention behaviour. Further study is needed to produce more accurate models. This can be done by considering additional factors such as customer transactions and demographics, in conjunction with customer complaint behaviour.

Second, "intention to stay" is often used as a proxy for customer retention (Fornell, Johnson, Anderson, Cha, & Bryant, 1996; Zeithaml, Berry, & Parasuraman, 1996). However, a repurchase intention is not necessarily an accurate predictor of defection (Bolton, 1998; Bolton & Lemon, 1999; Mittal & Kamakura, 2001). Similarly, survey

[3] The information is sourced from Gartner Inc. industry report subscribed by Firm A. The report is part of copyrighted Industry Report & Forecasts Series dated January 2010.

measures of "satisfaction" have also proven to be unreliable at times, with even "satisfied" customers eventually defecting (Chandrashekaran et al., 2007).

Third, a current hindrance to understanding customer retention is the lack of research conducted in the context of *non-contractual customer-firm relationships using a longitudinal data analysis approach*. Certainly, there are specific stages in an individual customer-firm relationship. A contractual environment is characterised by the existence of contracts between the firm and customer. It is usually focused on relational activities which heighten the importance of building and sustaining positive customer-firm relationship with the goal of increasing customer loyalty and commitment (Roos et al., 2004). In most cases the date of defection is clearly known since it matches up with the contract cancellation date. In a non-contractual environment the emphasis is mainly on transaction and usually customers have low degree of intention to stay with the firm (Kumar, Bohling, & Ladda, 2003). This implies that customers with a non-contractual relationship has low affinity and can change their pattern of purchases without informing the firm (Bolton, 1998; Danaher, 2002).

While, the change from a non-contractual to a contractual relationship is a continuous process this study is designed to examine non-contractual customer-firm relationship. A distinction between contractual and non-contractual is fundamental, because it is completely inappropriate to apply a model developed for a contractual setting in a non-contractual setting (Reinartz & Kumar, 2000).

Most existing research on subscription-based services has been conducted within contractual settings (Ascarza & Hardie, 2013; Bonfrer, Knox, Eliashberg, & Chiang, 2010; Dover & Murthi, 2006; Schweidel, Fader, & Bradlow, 2008). Furthermore, in contractual settings, customers are less likely to leave to avoid breaking a contract and paying an exit fee. Since a firm is able to observe the entry and exit dates of a customer, they can take proactive measures to address defection issues. Modelling retention for customers in a non-contractual relationship is critical because defection may be unobserved. Therefore, it is more important to study customer retention in a non-contractual setting.

1.9 Gaps in Current Models

There are limited studies on customer retention that uses actual customer complaint data. Within a service context, most research has focused on using traditional factors, which are identified as influencing retention, as a means of understanding retention behaviour. While helpful, is an oversimplification; research is needed to investigate inter-temporal data (Donkers, Franses, & Verhoef, 2003). Also, most research models are restricted by only using parametric and non-parametric approaches (Kamakura et al., 2005).

Further, current research is limited to understanding defection behaviour at aggregate levels and neglects differences that exist between individuals. Customers at risk of defection must be identified and individual customer-level views must be generated (Gupta et al., 2006; Gupta & Zeithaml, 2006). Additionally, current research on retention focuses on cross-sectional models. While it is necessary to understand the process that leads to customer retention, cross-sectional models are not sufficient to measure change over periods of time. Customer-firm interactions cannot be adequately studied with one-time snapshots of data; however, a longitudinal approach allows retention to be understood over a period of time.

1.10 Research Aims and Research Questions

Firms often focus more on acquisition than retention. Payne and Frow (2013) discuss two studies, which show approximately twice as much of the marketing budget is allocated to acquiring new customers, compared to retaining existing customers. A lack of attention to retention is due in part to a lack of adequate information on reasons for defection. According to Payne and Frow (2013) customer's weapons are a diminished sense of loyalty, resulting in a greater propensity to switch to organisations that may provide better service. The current global economic situation has increased cost-consciousness and this is reflected in customer attitudes towards pricing. As customers are becoming smarter buyers, firms that can provide uninterrupted services have a great advantage.

Given the motivation for this research and gaps in the current body of knowledge, this study aims to examine customer retention behaviour using customer complaint behaviour. This study uses longitudinal data over a period of three years to capture simultaneous effects of several explanatory variables using a dynamic statistical approach to analyse customers subscribing to products with two turning points in the markets, namely, the decline and growth stages of the product life cycle.

The following research questions are addressed in this study:

> **Research Question 1:** Can customer complaining behaviour explain customer retention behaviour?
>
> **Research Question 2:** Does the number of complaints explain customer retention behaviour?

Research Question 3: How does “recency” of complaints impact customer retention behaviour?

Research Question 4: Does the severity of complaints have greater effect on customer retention behaviour?

Research Question 5: How does the length of recovery influence customer retention behaviour?

Research Question 6: Does communication with customer support help or damage customer retention?

Research Question 7: Does length of relationship and usage explain customer retention behaviour over time?

Research Question 8: Are customer retention rates different depending on customer demographic characteristics?

1.11 Context of the Research

The confidential data used in this research has been obtained from a large telecommunications provider in Southeast Asia. This provider offers voice, Internet, high-speed broadband, and other services (prepaid calling cards and value added services) on a subscription basis to its customers. The analysis is based on a sample size of 15,000 to 20,000 active customers, randomly selected from the customer bases of two different products: voice and Internet. Randomisation of the sample over 36 months allows for customer heterogeneity across different geographic locations.

1.12 Modelling Choice

This research uses survival analysis in its model computation. *Survival analysis* describes the class of statistical methods for studying the occurrence and timing of events (Allison, 2010). Research suggests that this is the best method to operationalise this study because it factors in "censoring"[4] and the ability to manage time-varying explanatory variables. Not addressing censoring will result in biased, unreliable, under, or overestimated parameters (Nitzan & Libai, 2011). Logistic and "probit" models cannot account for continuous dependent variables (i.e., when a customer defects), but rather these models accounts for only dependent variables with two outcomes (i.e., either they defect, or do not defect). Findings show that customers do not behave uniformly, so unobserved heterogeneity must be factored into the defection model. Further, a more conventional methodology such as regression analysis would only explain the likelihood of an event occurring, without explaining when it might occur. This is the reason why survival analysis is a more appropriate methodology. A survival analysis model takes into account the risk of the event occurring over a period of time. If the customer has a high "hazard rate" (Allison, 2010; Helsen & Schmittlein, 1993; Singer & Willett, 2003), it means survival time in a given period is likely to be low. Inversely, if the hazard rate is low, survival time is likely to be high.

[4] Due to the use of actual behavioural data in this study, the dataset contains observations that are censored. Observations are considered censored when information about the survival time of customers in only partially known. That is, there could be customers who continued their relationship after the observation period. Under this circumstance the survival time is considered as partially known. It is possible to correct this type of biasness through appropriate methodology. Censoring is discussed in Chapter 4, Section 4.1.3.

1.13 Dependent Variable and Explanatory Variables

In this study a number of theoretical and empirical relationships between customer retention and a range of explanatory variables are investigated. The longitudinal data available for the study allows us to test the robustness of these results. *The dependent variable in our model is conceptualised as the hazard rate of an individual customer defection measured at time (t).* The goal of our research is not only to examine the effects of time on complaints, but also to assess the relationship of the survival time to the explanatory variables on the dependent variable.

The *explanatory variables* operationalised in this research are:

(i) *Behavioural measures on customer complaints that provides information regarding:*

(a) If a customer has complained
(b) How many complaints were made in the last month
(c) Recency of the complaint
(d) Severity of the complaint
(e) Length of recovery of the complaint

(ii) *The behavioural measures relating to relational exchange characteristics that provide customer relationship information are*:

(a) Number of communication encounters per month (excluding complaints)
(b) Usage level (measured by number of downloads per month and minutes)
(c) Length of relationship (measured in months)

(iii) *The customer characteristics examined are:*

(a) Age
(b) Gender
(c) Household income
(d) Household size

These variables were derived from the exhaustive search of literatures covered in number of studies and these additional variables encompass broader range of factors than previously covered in extant studies. This literature is discussed later in Chapter 2, Sub-section 2.7 and in Table 2.1.

1.14 Summary of Findings

This study uses the Cox Proportional Hazard model (PHM), a semi-parametric approach to survival analysis, to examine customer retention behaviour. Unlike parametric approaches, the Cox PHM model incorporates variables that are both time varying and time invariant. Further, random effects are introduced in the Cox PHM to address issue of repeated measurements and unobserved heterogeneity across customers.

Three models: *baseline*, *comprehensive* and *interactions* are estimated for voice and Internet subscribers using PHREG procedure in SAS[5].The models are evaluated using

[5] The analysis could be carried out using different statistical packages such as SPSS, SAS, Stata, R but SAS currently has the most comprehensive set of full-featured procedures for performing survival analysis (Allison, 2010).

Akaike's Information Criterion (AIC) or Schwarz-Bayes Criterion (SBC)[6] (Akaike, 1981; Schwarz, 1978) and the likelihood ratio test (LRT). Finally, in utilising the Cox PHM model, the assumption of proportionality is assessed, based on Schoenfeld residuals. The analysis procedure and all results (including an explanation of technical terms) are discussed in Chapter 5.

The final model for both Internet and voice services, the two products studied in this research, is derived from the interactions model and the results shows that retention behaviour of customers subscribing to voice and Internet are similar except for explanatory variable, usage. However, of the two products, the study found that generally Internet customers have lower defection rates compared with voice customers. The study also found:

Factors that influence customer retention behaviour based on *increased risk of defection* include:

a. Complainants versus non complainants
b. Length of recovery
c. Age
d. Gender
e. Household income

[6] AIC and SBC holds the same interpretation in terms of model evaluation. However, a smaller value of either AIC or SBC indicates stronger evidence for one model over the other. The information on model evaluation is provided in Appendix F.

Factors that influence customer retention behaviour based on the *decreased risk of defection* include:

a. Number of complaints
b. Recency of complaints
c. Severity of complaints
d. Usage of Service
e. Length of relationship

In addition to identifying the relationship direction, the relationship strength is also identified in the study. The study results and their implications are discussed in detail in Chapters 5 and 6.

1.15 Contribution to Knowledge

1.15.1 *Theoretical Contributions*

This research is among a limited number of existing studies that incorporate customer complaint behaviour, rather than perceptual metrics, in the context of customer retention research. Studies such as Knox and Van Oest (2014) and Lariviere and Van den Poel (2005) do consider customer complaints. However, they do not address other factors, such as number of complaints, effects of "recency", and other customer complaining/non-complaining behaviours. In this study a more complete understanding of complaint behaviour has been achieved through use of a comprehensive dataset comprising details of actual consumer complaint behaviour. As a result of this research helps explain factors affecting retention by understanding the means of decreasing or increasing the *risk of defection*. Further, this research has contributed to the body of relationship and services marketing literature, specifically, customer retention studies through a refined understanding of both behavioural metrics and usage, length of relationship, and consumer characteristics.

In addition, this research study *compares customer retention behaviour of two products at different positions on the product life cycle* - the growth and decline stages. This analysis provides an understanding of those factors that are common across two products and those factors that vary between them. This aspect of market characteristics appears to have been neglected in previous customer retention studies. Therefore, important contribution of this research is to investigate the characteristics of customer retention behaviour of products at different life-cycle stages.

This research models the key drivers of customer retention at a micro level, and longitudinally, through the inclusion of customer complaints, service failure, and service recovery. Previous studies have generally relied on perceptual metrics which may not vary over time. This research uses “big data”, comprising longitudinal data files from existing client records to investigate customer retention.

This empirical study, using a single dynamic hazard model, allows for time varying explanatory variables from different categories of independent variables to assess the incidence and timing of the customer’s actual behaviour. The analysis permits identification of key factors influencing customer retention and the directionality of the relationship.

By better understanding customer retention behaviour, businesses can start to develop more effective strategies for reducing defection and improving retention rates. This research should assist firms with similar subscription-based services in their efforts to bolster retention. Finally, the model in this study can be applied to other firms to help improve targeting efforts by prioritizing customers with the highest risk of defection.

1.15.2 *Methodological Contributions*

A semi-parametric approach includes explanatory variables that are time variant and invariant. As noted by Kamakura et al. (2005), most research in retention studies focus on using non-parametric and parametric approaches. To gain deeper insight into customer retention, we conclude that studies need to address two important issues. First, they need to explore semi-parametric models, in order to consider both time varying and invariant variables. Second, studies need to use a wider range of explanatory variables as a means of providing a more comprehensive understanding of factors affecting customer retention.

This current study includes several new explanatory variables including various aspects of customer complaints to understand customer retention behaviour. Further, although parametric approaches are relatively easy to interpret, they are limited in incorporating marketing data that is intertemporal. As a consequence a semi-parametric approach[7] is adopted in this study. The methodology adopted in this study includes a large number of explanatory variables in order to avoid omitted variable bias that could lead to over- or under-estimating the parameter.

---

[7] The approach is further discussed in Chapter 4, Section 4.2.2.

1.15.3 *Managerial Contributions*

This research presents a number of managerial contributions relevant to the telecommunications sector as well as the service sector more broadly. Senior management can use this research to better understand complaints and how complaints impact the customer base and sales revenue.

It is no longer sufficient for a single person or department to handle customer issues. Complaint data must be analysed fully. It is important to improve current operations cross-functionally across the whole firm (Ryals & Knox, 2001). The long term success of the enterprises rests on how effectively customer complaint behaviour is handled. All employees and managers must appreciate the far reaching effects of complaint behaviour on the firm. Managers' key performance indicators should address the parameters described in our model (see Chapters 5 and 6).

Shareholders and the wider stakeholder community should also be aware a firm's practices with respect to complaint behaviour in order to ensure lasting trust in a firm. The ability to understand customer retention behaviour provides an opportunity for a firm to develop approaches for minimizing lost customers and improving methods for customer retention.

Firms have limited resources. By understanding the factors that influence customer defection, this research will indirectly assist firms in balancing the financial resources used in or acquisition and retention strategies.

Finally, the model in this study can help firms prioritize and target those customers that are at risk of defection with appropriate products and service. For example, the firm may develop a targeted and proactive communication campaign to address profitable "at-risk customers". Additionally, specific programs may be developed to increase usage of the existing products or services.

1.16 Public Policy Implications

There also some public policy implications. Government and regulators can use customer complaint information to develop customer charters that balance the interests of consumers and industry participants. That is, Government or regulators can help ensure that a customer has access to quality services and ensure that complaints are handled effectively. Some regulators may encourage and facilitate customer complaints and monitor the level of complaints received. Regulators may move to a situation where they help ensure that industry actors comply with charters that focus on customers' rights.

1.17 Structure of the Thesis

This thesis is organised into the following chapters:

Chapter 1 provides an introduction to the study. It discusses the background to the research, the aims of the study and the research questions investigated. It also includes an overview of the research, and a short discussion of the research methods utilised. A brief summary of the findings and contributions of the study are presented.

Chapter 2 undertakes a review of relevant literature relating to this study. This chapter seeks to highlight key academic work in customer retention and defection, especially within the context of business-to-customer marketing relationships. This chapter comprises three main sections. The first section traces links between factors that influence customer retention and customer defection within subscription-based services. The second section identifies common models used in exploring customer retention and defection. The third section explains the goals of this research, the gaps in existing research, and describes more fully the formulation of the research questions. This chapter also discusses the context of the study-the telecommunications sector.

Chapter 3 addresses measurement and a description of the data sets. It provides information relating to the data, the variable operationalisation, the sampling, estimation issues. The model is presented in its mathematical form.

Chapter 4 describes technical aspects of the methodology utilized in this study. This chapter builds on the research goals and questions outlined in previous chapters. It discusses several alternative methods that could have been used in this research and explains why *survival analysis* is the most appropriate methodology. The chapter then discusses three forms of survival analysis that enable analysis of the relationship of a set of explanatory variables including: non-parametric, parametric and semi-parametric approaches. The chapter then explains why one particular type of semi-parametric approach, the Cox Proportional Hazard Model, best support this study's research aims and the structure of the research data.

Chapter 5 reviews statistical characteristics of the data and the results of the model analysis. The chapter also explains the analysis and interprets results, summarising key findings and comparing them with existing literature. This chapter specifically addresses how the study informs each of the eight research questions outlined earlier in this chapter.

Chapter 6 concludes the thesis by summarising the research findings and discussing limitations of the study. The chapter considers theoretical and practical implications of the analysis and makes recommendations for future research.

CHAPTER 2: LITERATURE REVIEW

2.0 Introduction

The literature presented in this review, draws on research in relationship marketing, customer relationship management, services marketing, subscription-based services, customer retention and defection (including customer switching and churn), service quality, customer complaint, service failure, and service recovery. This chapter is structured as follows. First, the chapter discusses literature relating to the choice of explanatory variables for this research. It considers key characteristics of customer attention and the explanatory variables that have been previously identified in extant research studies. Second, it reviews the research context and presents conclusions regarding the need for specific explanatory variables to be studied further, because of existing gaps in current studies of customer retention.

2.1 Relationship Marketing and Customer Retention

The term relationship marketing appears to have been first introduced in 1983 in an article by Berry (1983). Relationship marketing focuses on improving customer retention and reducing customer defection. Berry emphasized that attracting new customers should be seen as an initial stage in a marketing process, which focuses on establishing a relationship, turning apathetic customers into loyal customers, and treating customers as important clients (Berry, 1995).

This initial work by Berry urged academic scholars to shift their research focus from a transactional to a relational approach (Eriksson & Vaghult, 2000; Morgan & Hunt, 1994).

From the managerial perspective, Reichheld and Sasser (1990) captured the attention of academics and managers by proposing that profits soar when firms lower customer defection rates. Some years later Reichheld (1996) argued that firms could increase profits from 25 to 85 per cent by reducing customer defection by five percentage points. Reichheld (1996) further stated that another five percentage points decline in defection doubles the length of relationship, increasing profits , in the context he studied, by an additional 75 per cent (i.e., from $300 to $525). Although there was no detailed scholarly evidence supplied within their published work, customer retention started to receive substantive interest and academic research on relationship marketing increased. Since that time there has been a general shift from marketing activities focused principally on customer acquisition to those balanced in terms of focusing on both customer retention and acquisition. Following Reichheld and Sasser's (1990) work, many studies confirmed a positive relationship between customer retention and profitability (Gupta et al., 2004; Hallowell, 1996; Payne, Storbacka, & Frow, 2008; Van den Poel & Larivière, 2004). Following this body of research, the ability to retain customers became recognised as a key competitive advantage (Kim & Yoon, 2004).

According to Harker (1999), Grönroos' definition of relationship marketing is the best of 26 different definitions that he identified: "Relationship marketing is to establish, maintain, and enhance relationships with customers and other partners, at a profit, so that the objectives of the parties involved are met. This is achieved by a mutual exchange and fulfilment of promises" (Grönroos, 1994).

Dwyer (1989) was among the early researchers to study the relationship between retention rates and customer lifetime value. He proposed a customer retention model that estimated lifetime value. The positive correlation between retention rates and customer lifetime value has been further supported by other researchers (Fader & Hardie, 2010; Gupta & Lehmann, 2003; Reichheld, 1996). Venkatesan and Kumar (2004) concluded that if a firm wishes to maximize customer lifetime value, it should allocate more of its budget to customer retention strategies rather than customer acquisition strategies. Indeed, relationship marketing strategies are integral to promoting customer retention, lifetime value, and profitability (Bolton, Lemon, & Verhoef, 2004; Ryals, 2005; Sheth & Parvatiyar, 1995; Verhoef, 2003). A key part of relationship marketing strategy involves examining which customers are at risk of leaving a firm (churn), and allocating resources to retain those customers (Shaffer & Zhang, 2002). For example, customers at risk of churn may be offered special incentives (e.g., discounts, gift cards, coupons) to improve their relationship with the firm and increase retention (Burez & Van den Poel, 2007).

Since customer retention is an important driver of profitability, it is important to capture the correct explanatory variables to increase accuracy in retention models. Models that fail to identify potential churners or falsely indicate the propensity to churn can lead to substantial financial losses. Keaveney and Parthasarathy (2001) underline the critical importance of identifying customers who are prone to leaving the firm - it is far more profitable to satisfy existing customers than to attract new customers. Heskett and Schlesinger (1994) also examined potentially vulnerable customer groups, including those experiencing low-quality services, new customers, and customers of firms with sharp competition.

The building of early warning models to identify such customers is a complex research field crucial to business success. While current literature provides several defection models (Ahn et al., 2006; Coussement & Van den Poel, 2008; Hadden, Tiwari, Roy, & Ruta, 2006; Hung, Yen, & Wang, 2006; Kim & Yoon, 2004; Van den Poel & Larivière, 2004), future research needs to develop models that integrate the most important dimensions of retention, including customer complaint behaviour customer transaction history, and customer demographics.

2.2 Customer Retention and Defection

Gupta and Zeithaml (2006) have defined customer retention as "the probability of a customer being 'alive,' or repeat buying from a firm". Ahmad and Buttle (2001) describe customer retention as "the mirror image of customer churn, where a high retention rate has the same significance as a low churn rate." This latter description suggests that in order to improve customer retention, it is necessary to fully understand customer defection and the reasons underlying defection behaviour. In this current research, the terms customer churn and customer defection are used interchangeably to refer to the failure to retain a customer.

Scholars vary in terms of in terms of how they define churn. Lemmens and Croux (2006) defined churn as "whether a current customer decides to take his or her business elsewhere." Adopting a similar position, Buckinx and Van den Poel (2005) consider a customer is a churner only if all relationships with the firm are terminated. By contrast, Owczarczuk (2010) define a "churner" as a customer who has been inactive for at least six weeks. Others, e.g. (Ahn, Han, & Lee, 2006; Brockett et al., 2008; Lariviere & Van den Poel, 2005), have attempted to draw fine distinctions between total and partial defection. In particular, there may be circumstances that customers do not renounce the services of a firm all of a sudden, but through a gradual shift, exhibiting *partial defection* behaviour. For instance, customer may choose to obtain some of the needed services from another firm. In case of partial defection, there is a probability that after a while, they will defect completely to competitor. Therefore, in the long run, *partial defection* may turn into *total defection*. Nevertheless, whether a customer terminates the relationship, or becomes inactive for an extended period of time, the firm is clearly experiencing difficulty in retaining the customer's business.

Churn or defection can further be classified as "uncontrollable" and "controllable." Braun and Schweidel (2011) define "uncontrollable churn" as churn resulting from reasons beyond the influence and control of the firm, such as death, relocation, etc. "Controllable churn" is due to reasons that can be influenced by the firm's actions, such as service failures, service encounter failures, service recovery failures, etc. Research has found that consumers tend to assign greater blame on a firm when they perceive failures as controllable (Laufer, 2002) . Therefore, it is critical for a firm to accurately identify those customers most prone to churn due to controllable factors and to devise incentives for retention (Braun & Schweidel, 2011). The current study addresses this issue by focusing on controllable churn by examining customer complaint behaviour and customer transaction characteristics that firms can shape and influence. It also includes customer demographics which are largely uncontrollable except to the extent to which a firm adopts segmentation and positioning strategies aimed at particular demographics.

2.3 Overview of Factors that influence Customer Defection

Within the marketing literature a number of existing studies examine various factors influencing customer defection in various contexts. One of the most common approaches to explaining customer defection and retention involves investigating quality and satisfaction as key factors.

For example, customer switching is influenced by the perception of quality in the banking industry (Rust & Zahorik, 1993), promotions in the telecommunications sector (Schweidel et al., 2008), dissatisfaction levels in the insurance industry (Crosby & Stephens, 1987); and service encounter failures in the retail sector (Kelley, Hoffman, & Davis, 1994). Churn increases when tangible dimensions, such as quality (Dabholkar & Walls, 1999), value (Bansal & Taylor, 1999), and price fall short of customer expectations. Perceived costs of switching also reduce customer defection (Ping, 1993). However, high switching costs may also anger customers and increase defection as a form of revenge (Haj-Salem & Chebat, 2014). Other attitudinal factors, such as alternative attractiveness (Jones, Mothersbaugh, & Beatty, 2000) and attitudes toward switching (Bansal & Taylor, 1999), also influence customer retention.

Although tangible aspects of service, including service quality, are important drivers of customer behaviour, they do not always lead to a change in customer intentions (Marshall & Shepherd, 1999). Keaveney (1995) revealed that even satisfied customers might abandon their relationship with a firm. Roberts, Varki, and Brodie (2003) also argued that relationship quality performs more accurately than service quality as an estimator of behavioural intention. Berry emphasizes that marketers striving to build "the strongest possible relationship" must extend their efforts beyond price competition (1995, p.240). Colgate, Tong, Lee, and Farley (2007) provide a more comprehensive understanding of the customer-firm relationship in retail banking services using data collected from banking customers. Along with price, these latter authors identified other factors that influence customer churn, including switching barriers, low perceived differentiation between available alternatives, emotional bonds with service providers, confidence in their service providers, service failures, and good service recovery.

Therefore, tangible factors as well as relational factors are important considerations for understanding customer retention. However, in certain circumstances, relational factors such as employee empathy can negatively influence customer retention. For example, Prentice (2014) found that casino employees recognizing regular players and providing individualized attention has a negative effect on player patronage because customers do not wish to be identified and perceive their attention as unnecessary or unwelcome. In this context, empathy has an unfavourable influence on retaining customers.

A number of explanations have been put forward regarding the antecedents of customer retention. Aydin, Özer, and Arasil (2005) suggest corporate image and reputation could influence customer loyalty to a firm. Lee, Lee, and Feick (2001) regard customer confidence, enterprise visualization, perceived value, and communication techniques as instrumental to evaluating customer loyalty. Kim and Yoon (2004) findings show that, in the mobile telecommunications industry, the likelihood of customer defection depends on the call quality, tariffs, handsets, and brand image, but also customer incomes and subscription duration. In an online context, trust towards a website, the brand, and the information available are important determinants of online customer retention (Toufaily, Ricard, & Perrien, 2013).

Other recent studies provide further insights. Wirtz, Xiao, Chiang, and Malhotra (2014) find that a firm's marketing activities aimed at attracting new customers (advertising and distribution) can also help retain current customers. Malhotra and Kubowicz Malhotra (2013) find that "soft lock-in", where friends share the same service provider, lowers customers' intent to switch, whereas "hard lock-in", where there is high switching costs, leads customers to increase their intent to switch. This indicates that there are many perspectives to consider in order to better understand how firms can retain customers and prevent defection.

2.4 Customer Complaint Behaviour and Customer Retention

There is an important stream of research that positions customer complaint behaviour as a crucial dimension of the customer-firm relationship and customer retention. Fornell and Wernerfelt (1987) conclude that a firm's resources are better spent on retaining existing customers through effective complaint management than on acquiring new ones. Fornell and Wernerfelt (1988) also show that managing complaints efficiently can enhance customer retention. They investigate the implications of complaint-handling procedures on buyer retention, highlighting the relative importance of the efficiency of marketing spending resources on keeping customers, rather than attracting new ones. This suggests that understanding customer complaint behaviour is essential for increasing customer retention.

Further, complaint management is an important functional area of relationship outcomes and relationship management. As identified by Stauss and Seidel (2004), the importance of conflict and complaint handling in customer relationship management is an under-researched area. Additionally, research in this area has mostly been restricted to using satisfaction-related variables as surrogates in examining relationship outcomes. One of the main reasons for this is limited access to behavioural measures. For instance, Nitzan and Libai (2011) incorporate change in terms of number of service records as a proxy to understanding retention behaviour. However, this may not be an accurate measure of understanding customer retention behaviour. Recent service quality research by Brock, Blut, Evanschitzky, and Kenning (2013), and Sharma, Marshall, Alan Reday, and Na (2010) investigates the consequences of complaining behaviour. However, these studies are limited by using perceptual metrics such as behavioural intention. As Braun and Schweidel (2011) suggest, including operations metrics provides additional insight to understanding customer retention and increases accuracy.

Keaveney (1995) was also one of the first researchers to study the effect of complaint behaviour. The model she developed, using the Critical Incident Technique reported by 526 survey participants in 45 different service industries, formed her basis for future research. Keaveney (1995) identified core service failure, service encounter failure, and response to service failure among other factors, such as pricing, inconvenience, competition, ethical issues, and involuntary switching, as important determinants of customers' intention to leave a firm. Resolving customer complaints is an integral part of overall service delivery (Parasuraman, Zeithaml, & Berry, 1988), improving the quality of the relationship between the customer and the firm (Negi & Ketema, 2010) and increasing satisfaction with and trust in the firm (Alvarez et al., 2010). A complaint that is not addressed properly may serve as the most accurate signal of a customer's intent to leave a firm. Thus, Ahn et al. (2006) recommend that firms monitor the effectiveness of complaint management strategies to decrease customer churn.

Customer complaint behaviour is defined by Landon (1980, p.336) as "an expression of dissatisfaction by individual consumers (or on a consumer's behalf) to a responsible party in either the distribution channel or a complaint handling agency." Recent research identifies other potentially interesting factors influencing customer complaint behaviour. One such factor is anonymity of complainants. Strong ties to service employees may deter customer complaints. However, when customers have strong ties with service employees, concealing complainant identity encourages complaint behaviour (Vikas Mittal, Huppertz, & Khare, 2008). Membership in loyalty reward programs also deters customer complaints (Schumann, Wünderlich, & Evanschitzky, 2014). Research indicates that members in loyalty reward programs may overlook or discount severe service failures.

Among various factors that affect customer complaint behaviour is customer dissatisfaction, making customer satisfaction an important variable in predicting customer complaint behaviour (Heung & Lam, 2003).This finding has been confirmed by earlier researchers Liu and McClure (2001), Williams, Drake, and Moran (1993) and McQuilken, Breth, and Shaw (2001). Sharma et al. (2010) conducted a multi-national investigation on this issue and found that greater dissatisfaction increased the number of complaints, regardless of culture. Thus, across all cultures, customer complaint behaviour is an important expression of customer dissatisfaction caused by service failures. Further investigation of customer complaint behaviour would provide a greater understanding of customer retention.

A service failure is defined as "a real or perceived service-related problem or where something has gone wrong when receiving a service" (Maxham, 2001). A service failure occurs when a firm is unable to meet the customer's expectations (Smith, Bolton, & Wagner, 1999b). Keaveney (1995) identified two distinct groups of service failures: service encounter failures, such as uncaring, impolite, or unresponsive employees; and, poor responses to service failures, such as reluctant, negative, and lack of responses from employees.

The literature identifies several important aspects of service failures that firms should consider. For example, perceptions of high controllability over a failure, where a customer believes a firm could have prevented the failure, result in more negative customer reactions, including low satisfaction and low return intent (Choi & Mattila, 2008). Research indicates that strong relationships do not necessarily provide firms with a safety cushion in the event of a service failure (Haj-Salem & Chebat, 2014). Strong relationships may even backfire, as customers might feel even more betrayed in response to a service failure. Moreover, customers with no emotional bond with a service firm simply discontinue patronizing the service firm without notifying anyone of the problem (Kaltcheva, Winsor, & Parasuraman, 2013).

It is apparent that service failure has alarmingly negative effects on customer trust and satisfaction, resulting in higher chances of customer defection (Cranage, 2004). Even if past experiences with a firm have been positive, negative experiences from a service failure stand out in memory, which become the most salient information retrieved when evaluating the firm or considering repurchase options (Kalamas, Laroche, & Makdessian, 2008).

Customer defection is also a likely result of service encounter failure where incorrect information is given to customers (Bogomolova & Romaniuk, 2009). Keaveney's (1995) critical incident study of 835 customer-switching behaviours in service industries demonstrated that 44 per cent ended their relationship with the provider because of core service failures. Researchers including Bolton (1998), Bolton, Kannan, and Bramlett (2000), Kim (2000) and Mozer, Wolniewicz, Grimes, Johnson, and Kaushansky (2000) have concluded that service failures accelerate a customer's decision to discontinue the service provider-customer relationship. This negative effect of service failure on retaining customers is the underlying reason why customer complaint behaviour is so critical in understanding customer retention. Customer complaint behaviour research provides firms an opportunity to identify and rectify service failures, preventing and reducing the probability of customer churn.

2.4.1 *Customer Complaint Behaviour and the Focus of this Study*

In this research, customer complaint behaviour is examined by exporting a number of factors. These are discussed below.

Complainants versus non-complainants (the response parameter[8]):

In this study, whether a customer has made a complaint is examined to understand the likelihood of customer defection. Evidence suggests that not all dissatisfied customers actually complain, but often simply choose to defect (Blodgett & Anderson, 2000;

[8] The term "response parameter" or "parameter estimates" have been applied in variety of modeling duration data (e.g., Helsen & Schmittlein, 1993; Jamal & Bucklin, 2006; Kumar & Luo). The use of this term is due to the specific modeling approach adopted in this study. The signs of response parameter can be interpreted as similar to those in a regression model. Within the context of customer retention, a positive response parameter denotes a negative effect on customer retention. A negative response parameter denotes a positive effect on customer retention (Allison, 2010; Jain & Vilcassim, 1991).

Chebat, Davidow, & Codjovi, 2005). Dissatisfied customers who choose not to complain are also more likely to defect (Tax & Brown, 1998). McQuilken et al. (2001) conducted a survey of 171 pay-TV users and found that customers who chose not to complain when dissatisfied were significantly more likely to switch providers than those who complained. Thus, customers who have both complained and received adequate responses were more likely to continue their relationships with a firm (Conlon & Murray, 1996; del Río-Lanza, Vázquez-Casielles, & Díaz-Martín, 2009). In this regard, complaining behaviour can be positive for a firm, providing a chance to rectify mistakes and retain customers.

> *Therefore, this response parameter is expected to positively relate with customer retention behaviour for complainants as firms have greater opportunity to resolve issues*.

In this study, other specific aspects of customer complaint behaviour, are investigated including: (i) the number of complaints made by a customer; (ii) the time since last complaint; and (iii) the severity of the complaint are examined to understand customer retention behaviour. A greater practical and theoretical insight can be gained by characterizing customer complaint behaviour using these factors. These factors are now discussed.

Number of complaints:

As highlighted earlier, there are several reasons to contact a customer call centre, including service quality, billing, renewals, etc. For this reason, managing complaint volume is important in addressing customer retention.

Colgate et al. (2007) found that customers complained more often about service failures prior to exiting a firm. Hadden et al. (2006) utilised regression techniques, decision trees and neural networks to find the most significant variables in customer churn among 202 customers of a service provider. The most significant variables in predicting customer defection were the numbers and types of complaints. Ahn et al. (2006) studied data from 10,000 customers of one of the leading mobile telecommunications service providers in South Korea. They found that number of complaints is positively related to the probability of defection. However, a study by Bolton (1998) shows that increased calls provide a greater opportunity for a firm to rectify complaints and this increase is negatively associated with customer defection. However, intuitively, and increased number of complaints could indicate customers having prolonged service quality issues and this factor may be associated with a higher likelihood of customer defection.

> *Therefore, it is likely that number of complaints is associated with customer retention but the direction of the response parameter for number of complaints is uncertain. That is, the response parameter could be negatively or positively related to customer retention.*

Time since last complaint (recency):

A further potentially important characteristic of customer complaint behaviour is the effect of "recency". For example, Bolton, Lemon, and Bramlett (2006) provided empirical evidence that recency, in terms of service experiences, weighs heavily in the context of renewal decisions. For example, a customer placing many calls to a customer call centre within six months is more likely to defect than if those calls are placed within 12 months. Studies also confirm that customers who complain about service failures twice within a short time span are likely to have lower satisfaction (Maxham & Netemeyer, 2002b; Seiders & Berry, 1998) and this could influence defection.

> *Thus, the response parameter for time since last complaint is expected to positively relate to customer retention. That is, a customer who has recently made a complaint is more likely to defect than one who complained earlier.*

Severity of complaint:

Service failures are inevitable in most firms and varying severities of service failures are an important characteristic of customer complaint behaviour. Customers lodge complaints for a variety of reasons, including poor service quality, technical, and billing related issues. However, not all service failures are perceived as equal. Weun, Beatty, and Jones (2004) found that the severity of a service failure has a significant effect on satisfaction and subsequent renewals. In contrast, Keiningham, Morgeson, Aksoy, and Williams (2014) found that major incidents for airlines, such as accidents, injuries, and fatalities, demonstrate a weaker relationship with future market share than minor incidents such as flight cancellations. Also, major incidents showed no significant relationship with consumer satisfaction, whereas minor incidents were strongly and negatively related to future consumer satisfaction.

> *Thus, it is likely that the severity of service failure is associated with customer retention but the direction of relationship of this explanatory variable is uncertain as previous findings are not conclusive*.

Length of recovery:

In examining the effect of customer complaint behaviour on customer churn, service recovery cannot be ignored. Many researchers have identified service recovery as a key marketing tool for restoring customer satisfaction and retaining customer relationships after service failure (e.g., Miller, Craighead, & Karwan, 2000; Smith & Bolton, 1998; Smith et al., 1999b). Customers can experience higher satisfaction than their previous rating of satisfaction, even after a service failure, if the service recovery is performed effectively (Smith & Bolton, 1998). Good service recovery can turn a service failure into an opportunity to forge a long-term relationship with customers (e.g., Hart, Heskett, & Sasser, 1990). However, one empirical study found that only 60 per cent of customers are sometimes satisfied with a firm's service recovery effort, while almost 20 per cent of customers claim that they are rarely satisfied with service recovery efforts (Lovelock, Walker, & Patterson, 2011).

Mostert and De Meyer (2010), drawing on several authors, define service recovery as "the actions taken by an organisation to correct service failures by reinstating customers' level of satisfaction and loyalty to ultimately retain these customers" (Grönroos, 1990; Miller et al., 2000). Poor service recovery could cause negative word-of-mouth (Lewis

& McCann, 2004) and loss of confidence in a firm, reducing customer retention. When customers complain, they expect service recovery in a timely manner. Research suggests that a well-documented complaints-handling process can have a positive effect on customer retention (Fornell & Wernerfelt, 1987). Johnston (2001) identified a strong correlation between a firm's complaints handling prowess and customer retention. Further, Knox and Van Oest (2014) evaluates actual complaints and firms' recovery efforts on customers' subsequent purchases. This study adds to our understanding of the relationship between service recovery following a complaint and a customer's likelihood to defect.

Employees play an especially important role in recovering service failures. Prior research reports that a proper recovery action taken by frontline employees can bring positive consequences to the organization such as positive word of mouth and customer satisfaction (Blodgett, Hill, & Tax, 1997; Smith & Bolton, 1998). Improving a customer's perceptions of justice during the service recovery process is crucial for creating greater satisfaction, whereas inadequate service recovery worsens the customer's already low levels of satisfaction after the service failure (del Río-Lanza et al., 2009). Moreover, through proper emotional displays, frontline employees can mitigate customers' negative emotions after a service failure and prevent the contagion of anger among other customers (Du, Fan, & Feng, 2014). Research also indicates that service recovery can be complemented with an apology and/or compensation to achieve greater outcomes. An immediate apology has a significant impact on customer satisfaction. However, after the second week following a service failure, the apology has no significant effect on customer satisfaction (Fang, Luo, & Jiang, 2013).

Compensation also has a significant effect on repurchase intention. A study by Gelbrich, Gäthke, and Grégoire (2015) found compensation worth 70-80 per cent of a loss was optimal in producing the greatest satisfaction An immediate monetary compensation has the greatest recovery effect in the case of monetary failure (Roschk & Gelbrich, 2014). Customers waiting for a monetary loss to be rectified remain upset and do not recommend a firm to others. Joireman, Grégoire, Devezer, and Tripp (2013) recommend combining apologies and compensation for reducing customer desire for revenge and increasing desire to reconcile. Empathic employees cannot make up for poor compensation or ineffective recovery processes (Brock et al., 2013).

Hence, the time taken to resolve a complaint is a good indicator of a firm's ability to handle complaints and may affect customer retention behaviour (Davidow, 2003). Recovery time has been identified as an important criterion of procedural justice in complaint satisfaction, thereby increasing customer satisfaction and reducing churn (Blodgett et al., 1997; Clemmer & Schneider, 1996; Tax & Brown, 1998). Kau and Loh (2006) also studied customer responses to service recovery efforts using a cross sectional study of 428 mobile phone users. They found that the time taken to respond to a complaint is an important determinant of customer satisfaction.

> *Therefore, the response parameter for length of recovery is expected to positively relate to customer retention if the firm is successful in its recovery efforts.*

2.5 Customer Transactions and Customer Retention

Customer complaint behaviour and service recovery are critical dimensions for understanding customer defection; however, a broader perspective is also required to enhance the comprehensiveness of the customer retention model. Thus, customer transactional explanatory variables are included as other potentially important dimensions. Specifically, the customer retention model aims to provide better understanding for firms to maintain and enhance their relationships with customers. In this study, we examine relational exchange characteristics between a customer and a firm, including: (i) the number of communication encounters; (ii) the length of relationship; and (iii) service usage.

Number of communication encounters (intensity of interactions):

The intensity of interactions is one of the basic dimensions of a customer's transactional journey (Lemke, Clark, & Wilson, 2011) and a key determinant of relationship continuity (Crosby, Evans, & Cowles, 1990; Nicholson, Compeau, & Sethi, 2001). Thus, the number of communication encounters, which denotes the intensity of interactions between a customer and a firm, needs to be considered in exploring the factors that influence customer defection. As Payne et al. (2008) explain, the encounter process could occur at any level, and be initiated by either the customer, the firm, or both parties. For instance, a customer may call to enquire about a firm's product offering or to request information relating to specific service changes. Likewise, the firm's communication effort and responsiveness to customer enquiries play a crucial role in shaping customer-firm relationship outcome.

> *The response parameter is expected to be positively related to customer retention, that is, increased communication between the customer and a firm leads to increased retention.*

Length of relationship with firm:

The length of relationship represents how long a customer remains with a firm. Therefore, length of relationship is viewed as a basic component of a customer's transactional journey. A number of studies have found that the duration of customer-firm relationships is a key factor positively impacting retention (Bolton, 1998; Schweidel et al., 2008; Wirtz et al., 2014). For example, an extended customer-firm relationship signifies a high level of underlying involvement and overall service utility. In contrast, a customer dissatisfied with service utility or quality is likely to terminate their relationship at an early stage. In the context of non-contractual service relationships, the operationalization of the length of relationship is essential because a formal contract defining terms of the customer-firm relationship is absent. Customers in this context are free to change firms without paying a switching fee.

> *Thus, the response parameter is expected to positively relate to customer retention. That is, extended customer-firm relationships result in higher customer retention.*

Service usage:

Service usage is also incorporated in examining the likelihood of customer defection in this study. Kelley and Thibaut (1978) argue that usage is one of the basic dimensions defining customer-firm relationships. Bolton and Lemon (1999) also studied usage patterns and concluded that decreasing customer usage levels may indicate eventual defection. Despite this, it could also be argued that a customer who is a heavy user may actively seek better alternatives and, in turn, be more likely to defect. This point has been substantiated by Reinartz and Kumar (2003) and Prins and Verhoef (2007), who show that purchase frequency has a U-shaped relationship with likelihood of defection.

Hence, the variable usage is expected to positively correlate with retention if it follows the linear relationship.

2.6 Customer Demographics and Customer Retention

Age, race, sex, economic status, education, income, and employment are just a few demographic characteristics defining a population (Keep & Schneider, 2010). Kamakura et al. (2005) emphasize that customer relationship management (CRM) software used by firms must thoroughly analyse demographics such as age, location, and gender in providing reasonable anti-defection policies. In online contexts, customer demographics are also important in explaining customer retention behaviour (Keaveney & Parthasarathy, 2001). In this research, several demographic variables including age, gender, household income, and household size, are used to strengthen the customer retention model.

Age of customer:

In existing literature, personal characteristics such as age (Baumann, Burton, & Elliott, 2005; Hammond, 1998), gender (Melnyk, Van Osselaer, & Bijmolt, 2009; Mittal & Kamakura, 2001), and income (Hallowell, 1996; Reinartz & Kumar, 2003) have been found to affect customer defection.

Specifically, studies by East, Hammond, and Gendall (2006), and Wright and Sparks (1999) reveal that older respondents tend to be less loyal than younger or middle aged respondents. Conversely, studies by Baumann et al. (2005) and Patterson (2007) show that older respondents tend to stay longer with their suppliers. The socio-emotional selective theory argues that older customers tend to have limited time and cognitive resources (Cole et al., 2008). Therefore, they are less likely to develop new relationships with an alternate service provider.

For this research, we expect older customers to be more loyal and less likely to defect than younger customers. Therefore, it is expected that the response parameter for older customers is positively related to customer retention.

Gender:

Another personal characteristic that affects customer defection is gender. Studies by Melnyk et al. (2009) and Mittal and Kamakura (2001) found that female customers tend to be more loyal and risk averse than male customers. Wirtz et al. (2014) found that females and subscribers aged 50 years and older are more likely to switch. However, other studies by Jamal and Bucklin (2006) and Patterson (2007) found no significant association between gender and customer defection or loyalty.

Thus, it is likely that gender is associated with customer retention but the direction of relationship of this variable is uncertain as previous findings are not conclusive. That is, the effect of gender on retention is equivocal, and depends heavily on context.

Household income:

Household income and size are also expected to influence customer churn (Hallowell, 1996; Reinartz & Kumar, 2003). Income level has been found to affect customer decision-making and defection. The existing research suggests that a high income level has been negatively correlated with customer defection (Keaveney & Parthasarathy, 2001; Reinartz & Kumar, 2003). The understanding is that high income earners are less sensitive to price changes, especially in low involvement purchases. Additionally, high income earners may perceive higher switching costs as requiring more time and effort than switching is worth.

> *Thus, it is expected that a high income level is positively associated with customer retention.*

Household size:

Studies confirm that household size can influence customer churn behaviour (Mittal & Kamakura, 2001; Narasimhan, 1984). Vakratsas (1998) confirm finds smaller household is more susceptible to deals than bigger household and more likely to defect. This finding was supported by Buckinx and Van den Poel (2005) on the moderating role of household size.

> *Thus, the variable household size is expected to be positively associated with customer retention as bigger number of household size would have lower likelihood to defection than smaller family size.*

Table 2.1 summarises the expected effect signs of parameter estimates[9] for key explanatory variables included in the customer retention model described in Chapter 3. Table 2.2, which follows the next section summarises important gaps in the customer retention literature, shows the restricted scope of previous studies and the much more extensive coverage of variables undertaken in this study when compared with existing research.

[9] The term "parameter estimates" is also referred to as β (beta coefficients). It reflects the change in the response associated with a one-unit change of the variables, all other variables being held constant.The "effect signs" indicates the direction of the relationship (Allison, 2010; Jain & Vilcassim, 1991).

Table 2.1: Expected Relationship and the Effect Signs of Key Explanatory Variables in Customer Retention

Factors	**Explanatory variable**	**Description**	**Expected effect sign**
Customer complaint behaviour and customer retention	Complainant versus non-complainants	The response parameter is expected to positively relate with customer retention behaviour for complainants as firms have greater opportunity to resolve issues.	Positive
	Number of complaints	The direction of the response parameter for number of complaints is uncertain. That is, the response parameter could be negatively or positively related to customer retention.	Positive or negative
	Time since last complaint	The response parameter is expected to positively relate with customer retention. That is, a customer who has recently made a complaint is more likely to defect than one who complained earlier.	Positive

Factors	Explanatory variable	Description	Expected effect sign
	Severity of failure	The direction of the relationship is uncertain because previous findings were not conclusive. That is, the response parameter could be negatively or positively related to customer retention.	Positive or negative
	Length of recovery	The parameter estimates for length of recovery is expected to positively relate with customer retention if the firm is successful in its recovery efforts.	Positive
Customer transactions and customer retention	Number of communication encounter	The response parameter is expected to be positively related to customer retention, that is, increased communication between the customer and a firm leads to increased retention.	Positive
	Length of relationship	The response parameter is expected to positively relate to customer retention. That is, extended customer-firm relationships result in higher customer retention.	Positive

Factors	Explanatory variable	Description	Expected effect sign
	Usage	The response parameter is expected to positively relate with customer retention if it follows a linear relationship.	Positive
Customer demographics characteristic and customer retention	Age	It is expected that the parameter estimate for older customers is positively related to customer retention.	Positive
	Gender	It is likely that gender is associated with customer retention but the direction of relationship of this variable is uncertain. That is, the effect of gender on retention is equivocal, and depends heavily on context.	Positive or negative
	Household income	It is expected that a high income level is positively associated with customer retention.	Positive
	Household size	It is expected that bigger household size is positively associated with customer retention.	Positive

2.7 Summary of the Important Gaps in the Literature on Customer Retention

Complaint management is an important but often neglected aspect of relationship marketing (Alvarez et al., 2010; Stauss and Seidel, 2004). However, more recently, studies on service relationships including those of Brock et al. (2013) and Sharma et al. (2010) have started to investigate the consequences of customer complaint behaviour. However, these studies are limited by their use of surveys and perceptual measures, such as repurchase intentions and satisfaction, as surrogate measures of behavioural retention. Perceptual measures of retention may not be an accurate surrogate for actual churn behaviour (Mittal & Kamakura, 2001). Consumer retention and defection are often "proxied" by an "intention to stay" (Fornell et al., 1996; Zeithaml et al., 1996).

For instance, the findings of Keaveney (1995) regarding the relationship between customer complaint behaviour and retention were based on customer intentions, since they did not utilise actual firm data, but rather asked customers to respond to a survey. Nitzan and Libai (2011) also incorporated change in the number of service records as a proxy to understand customer retention behaviour. However, it is uncertain whether a repurchase intention is an accurate predictor of defection (Bolton, 1998; Bolton & Lemon, 1999; Mittal & Kamakura, 2001). Similarly, survey measures of 'satisfaction' have also proven unreliable at times, with 'satisfied' customers eventually defecting (Chandrashekaran et al., 2007). As suggested by Braun and Schweidel (2011), including behavioural metrics provides additional insights to understanding customer retention.

Studies in customer retention have typically dealt with only a few specific factors instead of using an integrated, comprehensive model to examine customer retention behaviour.

Research including factors like customer transactions and demographics in conjunction with customer complaint behaviour is needed to produce a more accurate customer retention model. Ahn et al. (2006) investigated several factors, including service usage, customer dissatisfaction, and switching costs. However, the only variable used to understand customer complaint behaviour was "number of complaints". To study customer complaint behaviour in further detail, research should be conducted with a greater number of variables, such as frequency, "recency", severity of complaints, and resolution time, resulting in a deeper understanding of the effect of customer complaint behaviour on customer defection.

Further, most research in customer defection using customer complaint behaviour has analysed customer data over a very short period of time with relatively small sample sizes. For example, Ahn et al. (2006) utilised other three months of data in their investigation. Also, sample sizes in this area of research were small: 171 by McQuilken et al. (2001); 201 by Hadden et al. (2006); and, 270 by Sharma et al. (2010). Edvardsson and Roos (2001) surveyed 60 telecommunication users in Sweden, utilising a critical incident technique (CIT). They found that a high number of complaints led to a weakened relationship with the firm and increased the chance of defection. However, their small sample size of 60 may limit the findings of this study. Moreover, the authors used a CIT method to analyse the customer interviews. As stated earlier, retrieval bias may limit the accuracy of CIT data (Edvardsson & Roos, 2001).

Additionally, other research in customer defection using customer complaint behaviour analyses customer data over a very short period of time with relatively small sample sizes. For example, Ahn et al. (2006) utilised three months of data in their investigation. Also, sample sizes in other studies in this area of research are small: 171 in McQuilken et al. (2001); 201 in Hadden et al. (2006); and, 270 in Sharma et al. (2010). Edvardsson and Roos (2001) surveyed 60 telecommunication users in Sweden, utilising a critical incident technique (CIT). They found that a high number of complaints led to a weakened relationship with the firm increased the chance of defection. However, their small sample size of 60 may limit the findings of this study. Moreover, the authors used a CIT method to analyse the customer interviews. As stated earlier, retrieval bias may limit the accuracy of CIT data.

Another gap in the recent customer retention literature is a lack of research conducted into the context of *non-contractual* customer-firm relationships using a longitudinal data analysis approach. Customers under a non-contractual relationship are more susceptible to defection as they can exit without incurring switching fees. Most research within retention literature has been conducted within contractual settings (Ascarza & Hardie, 2013; Bonfrer et al., 2010; Dover & Murthi, 2006; Schweidel et al., 2008) and non-contractual studies are limited. In contractual settings, customers are less likely to defect to avoid breaking a contract and paying an exit fee. In addition, since the firm can observe the entry and exit dates of a customer, it is possible to proactively address customer defection issues. Modelling retention for customers in a non-contractual relationship is critical. Unobserved defection increases firm uncertainty. While it is more challenging to model customer defection in a non-contractual setting, it is crucial to an accurate defection model. A few exceptions exist. Reinartz and Kumar (2003) have examined a non-contractual setting in estimating customer lifetime duration. Likewise, Fader, Hardie, and Shang (2010) have modelled the frequency of transactions in non-contractual relationships, but at a discrete point in time. However, it is argued that the consumption process occurs continually while customer-firm transactions can occur at any specific point in time. Thus, it is important to predict customer defection in a non-contractual context using a longitudinal approach.

Table 2.2 summarises these important gaps in the customer retention literature and illustrates the restricted scope of previous studies and the much more extensive coverage of variables undertaken in this current study. As mentioned previously, current literature predicting customer churn using customer complaint behaviour has several limitations. This research looks to fill these gaps by utilising behavioural measures of a larger sample size, over a 36-month period, in a non-contractual setting. This extended investigation period allows for greater analysis of customer defection with greater accuracy. Also, although prior research has examined traditional drivers of customer retention, such as customer transactions and demographics, studies including the role of customer complaints are limited. Combining information on individual transactions, demographics, and complaints is likely to provide additional insights into managing customer retention. A more rigorous approach, including a larger number of explanatory variables, avoids omitted variable bias that could over- or underestimates the parameter coefficients.

In sum this research develops a more comprehensive understanding of customer defection by incorporating a more detailed analysis of customer complaint behaviour, customer transactions, and demographics. By incorporating finer dimensions of customer complaint behaviour this study not only provides a more accurate customer defection model, it should also permit practitioners to identify valuable information about systemic service-related problems and identify potential new opportunities.

Table 2.2: Summary of the Existing Literature on Customer Retention and Defection

Objectives	References	Dependent variable: Actual	Dependent variable: Perceptual	Customer-firm interaction: Customer complaints: Complaints	Customer complaints: Frequency	Customer complaints: Recency	Customer complaints: Severity	Customer complaints: Recovery	Customer transaction: Encounter	Customer transaction: Tenure	Customer transaction: Usage	Customer characteristic: Age	Customer characteristic: Gender	Customer characteristic: Household income	Customer characteristic: Household size	Market characteristic: Decline	Market characteristic: Growing	Modeling approach	Longitudinal
Defection/churn probability	Ahn et al. (2006)	x		x	x						x		x			x		Logistic regression	8 months
Renewal behaviour	Ascarza and Hardie (2013)	x								x	x							Hidden markov	48 months
Customer satisfaction	Bolton et al. (1998)	x			x					x	x					x		Proportional hazard	22 months
Repurchase intention	Bolton et al. (2000)		x								x							Logistic regression	12 months
Retention and service experience	Bolton et al. (2006)	x				x			x	x								Nested model	12 months
Defection/churn probability	Bonfer et al. (2007)	x									x		x				x	Arithmetic brownian motion	12 months
Defection/churn probability	Borle et al. (2008)		x								x							Hierarchical bayes	12 months
Defection/churn probability	Braun and Schwidel (2011)		x							x		x		x		x		Bayesian hierarchical competing risk	18 months
Defection/churn probability	Buckinx and Van den Poel (2005)	x								x	x				x	x		Logistic regression, neural networks and random forests	5 months
Defection/churn probability	Burez & Van den Poel (2007)	x								x	x					x	x	Logistic regression, random forecasts & markov chains	12 months
Complaining behaviour	Chebat et al. (2005)		x	x			x											Principal factor analysis	
Retention/Satisfaction	Cooil et al. (2007)		x								x	x		x				Regression	24 months
Defection prediction	Coussement and Van den Poel (2008)	x								x	x	x	x					Logistic regression and generalized additive model	3 months
Service recovery	Del Rio-Lanza et al. (2009)		x				x	x										Structural equation modeling	
Customer lifetime value	Fader et al. (2005)	x									x							Pareto/NBD	18 months
Service recovery	Fang et al. (2013)	x					x	x									x	Vector Autoregressive (VAR)	10 months
Customer churn	Jamal & Bucklin (2006)	x							x	x	x	x	x					Parametric hazard models	12 months
Complaints	Kaltcheva et al. (2013)		x	x														structural equations model	
Complaints	Kau & Loh (2006)		x	x				x				x	x					Factor analysis	
Switching behaviour	Keaveney and Parthasarathy (2001)		x								x	x	x	x				Discriminant function analysis	3 months

Objectives	References	Dependent variable: Actual	Dependent variable: Perceptual	Customer-firm interaction — Customer complaints: Complaints	Frequency	Recency	Severity	Recovery	Customer transaction: Encounter	Tenure	Usage	Customer characteristic: Age	Gender	Household income	Household size	Market characteristic: Decline	Growing	Modeling approach	Longitudinal
Service failure	Keningham et al. (2014)	x		x			x											Partial least squares path	12 months
Churn and loyalty	Kim and Yoon (2004)		x							x		x	x	x		x		Binomial logit	
Complaints and churn	Knox and Van Oest (2014)	x		x				x			x						x	BG/BND	30 months
Post complaint and churn	Larivière and Van den Poel (2005)	x		x		x	x											Proportional hazard	12 months
Churn prediction	Lemmens & Croux (2006)	x									x	x		x		x		Logistic regression & stochastic gradient boosting	2 months
Switching behaviour	Malhotra and Malhotra (2013)		x							x		x	x					Exploratory factor analysis	
Service recovery and intentions	Maxham (2001)		x	x				x										Manova	
Service failure and recovery	Maxham & Netemeyer (2002)		x	x			x	x										Mancova	20 months
Customer loyalty	Melnyk et al. (2009)		x									x	x					Linear regression	
Satisfaction and repurchase intention	Mittal et al. (2001)		x									x	x	x	x			Anova, regression	
Retention probability	Nitzin & Libai (2011)	x		x						x	x	x	x	x		x		Proportional hazard	12 months
Loyalty and demographics	Patterson (2007)		x									x	x					Anova	
Customer lifetime	Reinartz and Kumar (2003)	x								x	x	x		x				Proportional hazard	36 months
Service failure	Schumann et al. (2014)		x				x				x	x	x					Factor analysis	
Retention	Schweidel et al. (2008)	x								x	x							Proportional hazard	24 months
Complaints	Sharma et al. (2010)		x	x														EFA	
Service failure and recovery	Smith and Bolton (1998)		x				x	x										Ordinary least square	
Satisfaction and service failure	Smith et al. (1998)		x				x	x				x	x					Nested model joint F-tests	
Customer attrition	Van den Poel and Larivière (2004)	x		x						x	x	x	x					Proportional hazard	12 months
Retention-complaints, transaction and demographics	**This study**	**x**		**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**x**	**Proportional hazard**	**36 months**

2.8 Research Context: Telecommunications Sector

In examining the effects of customer complaint behaviour, customer transactions and demographics on customer churn, this thesis focuses on the context of non-contractual, subscription-based telecommunications services. Subscription-based services such as telecommunications, utilities, and banking all operate under conditions of intense competition and market maturity, increasing the importance of building and sustaining positive customer-firm relationships, increasing customer loyalty, and decreasing defection.

Many subscription-based services offer associated services (such as equipment rental for connection to a public utility system) for a base fee and incrementally charge each transaction beyond an established minimum. For example, telephone and electricity services typically charge a base fee for connection and equipment rental, in addition to incremental charges for consumption beyond an established minimum. In subscription-based industries, defection translates to a failure to re-subscribe to a service. This has been examined by Coussement and Van den Poel (2008) in the context of a newspaper subscription service, and Burez and Van den Poel (2007) with a pay-TV subscription service.

The topic of complaint management warrants placing greater attention on customer retention in the telecommunications sector, as defection rates are alarmingly high. Within the wireless telecommunications industry, the monthly estimated rate of defection is about 2.2 per cent; which means that a firm loses about 27 per cent of its customer base every year (Wei & Chiu, 2002). Customer churn is an important outcome for firms, especially for mobile telecommunications service firms where the annual churn rate ranges from 20 to 40 per cent (Ahn et al., 2006).

Over the last two decades, many countries have deregulated telecommunications and utility services causing competition to surge in increasingly mature markets. The telecommunications industry can be described as highly competitive, mature, and plagued by high rates of churn (Ahn et al., 2006). Governments in many countries now encourage wireless portability, thus removing a major obstacle to changing mobile carriers. Additionally, consumers in many metropolitan areas have a choice of many mobile phone carriers.

For telecommunications customers, dropped calls and other service failures are extremely damaging and easily trigger their defection (Ahn et al., 2006; Chandrashekaran et al., 2007). For this reason, proper complaint response is critical to retaining customers. For high-volume businesses, unsatisfactory service is especially detrimental (Alvarez et al., 2010; Chandrashekaran et al., 2007).

The preceding discussion on market characteristics across different countries poses an interesting question: Do the effects of customer complaint behaviour, customer transactions, and demographics differ among markets with varying growth characteristics? Evidence suggests that customer retention is more critical for mature markets (Athanassopoulos, 2000; Jones et al., 2000).

In this study we examine two product markets, including a voice call market and an Internet market, which display respectively the characteristics of both market growth and market decline in terms of their stages in the product life cycle. In contrast to the voice call market, the Internet market is growing, which may lead to different effect sizes for the antecedents of customer churn, compared with those in the declining voice call market. This thesis examines both of these product markets to illustrate differing antecedents of customer churn, which may be influenced by growth market characteristics.

2.9 Methodological Approach

Much of the existing empirical literature in marketing has traditionally favoured parametric models that are easy to interpret. Such parametric models include probit or logistic regression, zero-inflated Poisson models, or parametric hazard specifications. However, defection studies need new approaches (Kamakura et al., 2005). New approaches are especially necessary in the telecommunications industry as most defection studies are focused on building prediction models rather than explanatory models. For example, in a study of customer churn in Taiwan, Hung et al. (2006) showed that neural networks are more accurate estimators than decision trees. Wei and Chiu (2002) investigated Taiwan's telecommunications industry and built a prediction model to estimate defection rates with subscriber contractual information and call pattern changes. Nath and Behara (2003) examined customer churn with the help of a Naïve-Bayes algorithm in the US mobile telecommunications sector.

Research in this field should attempt to modify flexible, semi-parametric models to handle the unique facets of marketing data. For example, researchers might use semi-parametric models to handle source heterogeneity that is unobserved in longitudinal data settings that comprise multiple customers' records (Bolton et al., 2004; Kamakura et al., 2005). To handle large numbers of explanatory variables or rare-event data, we identify that further research is required using semi-parametric methods.

2.10 Summary

The literature outlined in this chapter, provides a basis to further refine understanding of customer retention behaviour. Specifically, it is apparent from the literature review that limited attention is given on understanding customer retention using customer complaining information. The literature review also indicate that no studies have compared retention behaviour of customers subscribing to two products in the product life stages: declining and growing. The study aims to bring together three sources of data: customer complaints, customer transactions and customer characteristics to develop a more coherent understanding of customer retention behaviour.

The approach to measurement and data description is now outlined in the following chapter. In particular reason for choosing subscription based services and telecommunication are further discussed. Clearly, attributes of the services should influence the selection of customer retention model. The operationalisation of the dependent and the explanatory variables are further delineated.

CHAPTER 3: MEASUREMENT AND DATA DESCRIPTION

3.0 Introduction

The aim of this research is to understand the effects of various explanatory variables influencing customer retention behaviour. To support this aim, this chapter provides an overview of data requirements and research context, and describes properties of data and measurements in detail. First, it provides industry and firm requirements of data descriptions. Second, it introduces the dataset and describes properties of the data required for our research. Third, it describes available data sources and provides an overview of the dataset. Fourth, operationalisation of the dependent and the explanatory variables are presented, followed by the data exploration analysis procedure and variable selection techniques. Finally, the chapter concludes with a summary of the descriptive statistics.

3.1 Context

Research questions posed in the previous chapter can only be answered using a particular type of data. This data must come from an industry in the services sector that is both subscription-based and provides a range of products. The model developed is then estimated and tested using data from an incumbent subscription-based firm offering a variety of telecommunications services in emerging markets. The industry type was selected based on the reasons outlined below.

3.1.1 *Services*

The services sector provides fertile ground for modelling customer retention. As noted by Jain and Singh (2002) selling associated services and retaining customers has become more critical than acquiring new customers. Additionally, the salient characteristics of services, such as intangibility, heterogeneity, and the variability in consumption services, evoke key concepts of service quality and service experience. Because the model requires information on customer touch points, usage, and other behavioural data to understand customer retention behaviour, there is opportunity to empirically test and validate our research questions.

3.1.2 *Subscription-based Services*

Subscription-based services represent a suitable industry context for several reasons. First, subscription-based services provide a unique opportunity to observe and model customer-firm relationships over a period of time. As noted by Berry (1995), repeated contact between customers and service providers facilitates retention. In this instance, customers who subscribe to services offered by a firm may also choose to defect. This is true for many customers whose contracts expire and continue on a non-contractual basis. These attitudinal behaviours are both required in the model estimation and offer insights to what impedes customer retention. Second, subscription-based service firms know the identities of current customers and their general opinions regarding the services they offer. Hence, they provide a suitable environment to collect and model this type of data. Finally, subscription-based services are found in almost all areas of economic activity and account for 80 per cent of total services worldwide (Danaher, 2002; Lovelock, 1983).

This prevalence has proven instrumental in shaping research for the subscription-based model and is critical to understanding customer retention behaviour.

3.1.3 *The Telecommunications Firm*

The research data used in this study is from a telecommunications firm in Southeast Asia (hereinafter referred to as 'the firm' or 'Firm A') offering subscription-based voice, Internet, high-speed broadband, and other value added services. The firm has provided services since 1987 and has faced increased competition since the 1990s. There are four main competitors addressing this US$9.8 billion market valued for 2012[10]. The market share of Firm A compared to its competitors is provided in Table 3.1 for the period between 2010 and 2012.

[10] The information is sourced from Business Monitor International (BMI) telecommunication subscribed by Firm A. The report is part of BMI's copyrighted Industry Report & Forecasts Series dated January 2013.

Table 3.1: Market Share

Firm\Year	2010	2011	2012	Market/Products
Firm A	24%	22%	20%	Voice, Internet, high-speed broadband, pay TV
Competitor 1	21%	23%	23%	Mobile, fixed broadband, Pay TV
Competitor 2	19%	21%	22%	Mobile, VoIP, broadband
Competitor 3	16%	14%	18%	Mobile, VoIP, broadband
Competitor 4[11]	13%	12%	12%	Direct to home TV
Others (ISP's)[12]	7%	8%	5%	Prepaid Internet & VoIP

Source: Firm A's Financial Report, 2013 and International Data Corporation (IDC) Research Report 2013 and it is a copyrighted report.

Table 3.1 above shows that the firm operates in a competitive market and that customer defection is likely to occur as a switch to competitors rather than an abandoning of services. This provides a good context in which to model customer retention. When threatened by competitors in a monopoly market environment, the current focus on acquisition should shift to a focus on retention (Syam & Hess, 2006).

Internet and telephone services are becoming a larger, more integral part of daily life for many people. Because many consider these services essential, reliability is critical. Because necessity and reliability drive customer expectation, unsatisfied customers are

[11] Competitor 4 could provide Internet services via their fibre facilities and is an indirect competitor to Firm A.

[12] Internet Service Provider referred to as ISP's. The concentration of the service offering is limited to the city.

pushed to defect rather than to abandon services. Further, as noted earlier, the firm offers different types of subscription-based services including voice, Internet, and high-speed data. Voice-related services have been on the market since the pre-competition era. Using the term commonly applied in product life cycle (PLC) literature, voice services are in the "decline stage". Decline is the period of steadily decreasing sales before a product's demise (Golder & Tellis, 2004). For Firm A, this product has experienced no growth, combined with a high defection rate of 29 and 35 per cent in 2011 and 2012 respectively. To manage dwindling voice services customers, the firm does not impose any contractual obligations as they could be viewed as switching barriers or costs. This is an important element as the duration of the relationship is unobserved and this lack of switching barrier could be a key driver of defection.

The firm's Internet services were introduced post-competition. According to PLC literature, this service is in the "growth stage" between a new product's take-off and its maturity and decline stages (Golder & Tellis, 2004). While the firm has been struggling with revenue growth in the face of competition, defection rates are moderate, ranging from 15 to 17 per cent in 2011 and 2012 respectively. To manage increasing competition, the firm initially created a switching barrier which requires customers to have at least a minimum contract of 12 months. Later, they provided options of an increased subscription fee for customers with no contract, and a lowered fee for those with a minimum contract. Consequently, this study can examine customer retention behaviour for two products at two distinct points in PLC (e.g. the decline stage and the growth stage).

Regarding data availability, the firm collects and stores various forms of customer information, including operational, behavioural, and demographic data. This information is used in this study to explore factors affecting customer retention behaviour and for estimation purposes in the retention model.

3.2 Data Requirements

As noted earlier, the aim of this research is to understand the effect of various explanatory variables on customer retention behaviour across time, and in turn, to estimate customers' survival probabilities considering these variables. To gain insights into these factors, we now consider specific salient features of the data.

First, longitudinal data is collected at multiple points in time (each customer is measured repeatedly) to capture differences in customer behaviour and to understand customer retention. Cross-sectional data leads to limited understanding of customer retention, as behaviour can only be observed at a specific point in time (Blattberg et al., 2008). Further, the data must contain some defined interval of time for each individual customer. As noted by Jenkins (1995), the defined time measures against events occurring during that interval. Therefore, it is necessary to obtain information that captures the start date, or origin time, of a customer-firm relationship.

Second, as noted in the introduction, our research aims to include existing traditional drivers of customer retention, such as length of customer-firm relationship and transactions reflecting customer utility. Thus, individual customer information with an extended transaction history is required. Further, important information within the data concerns the type of customer-firm relationship; specifically, whether the customer has a contractual or non-contractual relationship. The contract status of a customer is pivotal in considering customer retention, behaviour because non-contractual relationships illustrate entry dates but not exit dates (Fader & Hardie, 2010). Consequently, non-contractual customers are more likely to defect, as they may do so without incurring a switching fee.

Third, another research goal is to understand customer retention behaviour using key factors identified with respect to customer complaints. To address this issue, observable and temporal information on customer touch points is necessary. Information on the number and severity of complaints can impact retention behaviour. Incorporating and applying such data allows for a more complete view of customer retention behaviour, which current marketing literature does not provide (Kamakura et al., 2005). Following the data requirements, the sub-section below describes three sets of databases provided by the firm for this research.

3.3 Databases

The following sub-section describes the structure and characteristics of the three databases recorded by the firm. It is important to note that the firm has provided subscription-based services since the 1980s. Since services began, the firm has progressively invested in their technology platform. This facilitates collection of multiple data sources including service, demographic, and operational data. The dataset available contains a wealth of customer level information that permits a detailed investigation of customer retention behaviour.

Datasets for two products were extracted from three firm databases. Although the types of data collected may differ across industries, most subscription-based firms operating in relatively mature markets, such as telecommunications, financial services, retail, and media, use customer information files. A comprehensive set of variables was identified from these data bases drawing on three categories proposed by Blattberg et al. (2008). These are explained briefly in the section below and in more detail in Section 3.5.

(a) *Customer identification data*: providing specific customer identification and classification. Some of this information includes unique customer identity (ID), home and business addresses, date of birth[13], and contact numbers. For example, while a unique ID may not be useful for making direct inferences on customer retention behaviour, it would aid in consolidating multiple databases. Furthermore, once this unique ID is assigned to a particular customer, it is used for all future interactions, even when a customer defects and returns to the firm. Although there were other important variables, such as address, that could be used to create another variable location, this information should be removed for non-disclosure purposes.

(b) *Demographic information*: included in this dataset are gender, household income, household size, occupation, number of children, marital status, and race. The initial retention model uses customer characteristics as a way to understand customer retention behaviour (Mittal & Kamakura, 2001; Reinartz & Kumar, 2003).

[13] The date of birth provided is converted to "age" and is usable as one of the demographic variables in the baseline model.

Demographic characteristics are useful to profile customers with high or low survival probability.

(c) *Transaction data*: refers to customer purchase behaviour stemming from specific use of a product or service. The transaction data was collected from various channels including customer interface, fulfilment, and billing. This type of data is generated internally and captures all aspects of customer-firm interactions. Transaction data includes customer activation and end date; payment; usage (in minutes and downloads); customer segmentation code; type of product packages; payment method.

In addition to the above category identified by Blattberg et al. (2008), this study extends the information to customer complaints as follows:

(d) Customer complaint data is also included in the transaction records. Complaint information includes: number of complaints; type of complaints; severity of complaints; recovery effort.

These data reveal the following information, which is helpful for understanding customer retention behaviour:

(a) Actual behaviour of dissatisfied customers determined by complaint behaviour and its consequences on future customer-firm relationships.
(b) Duration of customer-firm relationship on an individual level.
(c) Actual usage measures capturing changes over time for each customer, signifying intensity and depth of a customer-firm relationship.

(d) The decision to remain with or defect from the firm, reflected in customer commitment.

3.4 Dataset Description

Data from random samples of 80,000 active customers were extracted from firm's three databases. Because these were extracted from the firm's data warehouse, a simple random probability approach was observed. That means each customer within the population has an equal chance of inclusion (Coussement, 2014). Although the observation period is from 1st January 2010 to 31st December 2012, some of the customers were with the firm prior to the observation period.

As a result, this dataset may contain both left and right censoring[14] (Allison, 2010; Helsen & Schmittlein, 1993) which can influence the estimation of customer survival probability. This occurrence is widespread when using duration type data such as this. There are many types of censoring and the most common in subscription based services are left and right censoring (Allison, 2010). Right censoring occurs when a customer is retained by the firm at the end of the study observation period. Hence, it is unknown if the customer continued the patronage beyond the observation period. Left censoring arises when the customers start date with the firm was not observed within the study observation period. For example, left censored customers may have been subscribers to the firm's services prior to 1 January, 2010 and in this situation, the start date was not observed. However, this is

[14] Due to the use of actual behavioural data in this study, the dataset contains observations that are censored. Observations are considered censored when information about the survival time of customers in only partially known. That is, there could be customers who continued their relationship after the observation period. Under this circumstance the survival time is considered as partially known. It is possible to correct this type of biasness through appropriate methodology. Censoring is discussed in Chapter 4, Section 4.1.3.

not a problem in the dataset utilised because of the spectrum of subscription lengths by the customers.

Measures were taken to both manage and account for these biases and ensure a robust methodology. These measures are explained in detail in Chapter 4, which discusses the methodology in this study.

Additionally, the data contains customer information files for two products at different phases of the PLC. Examining similarities of customer retention between the two product categories at different PLC stages is fundamental to the research aims of this study. Therefore, to isolate the effects of product density on customer retention, only customers who purchase single services from the firm will be included in the model estimation. This is important because customers subscribing to more than one service may have higher switching costs, thus, affecting customer retention behaviour (Burnham, Frels, & Mahajan, 2003). The following section describes the dependent variables and the explanatory variables included in the model.

3.5 Variables Operationalisation

The firm provided activity data for each of the customers in the sample. From this detailed data, several key explanatory variables were utilised to help understand the relationship outcome (i.e., customer retention behaviour). This section presents an explanation of the conceptualization of the dependent and explanatory variables used in the customer retention model.

3.5.1 *Dependent Variable*

The dependent variable in this model comprise of two components: (i) status and (ii) duration. It is conceptualized as the hazard rate[15] of an individual customer defection measured as "time (t)" (Allison, 2010; Cox, 1972). The goal of the research is to not only examine the effects of time, but also to assess the relationship of survival time and explanatory variables on the dependent variable.

Following Nitzan and Libai (2011), the dependent variables receive two indicators for each customer. These indicators represent status and duration. The *status variable* differentiates between customers who have and have not defected. So, if a customer has defected before the observation period, the status variable has a value of "1". However, if a customer remains within the firm until the end of the observation period, the status variable has a value of "0". The *duration variable* differentiates customers who have and have not defected, measured between the start date of 1 January 2010 and end date of 31 December 2012. The model is estimated using a continuous time process because the precise time of events are known (Singer & Willett, 2003).

In a competitive market place it is assumed that customers could have defected to another market place or service provider. This assumption is logical, as customers in the retention model have non-contractual relationships with the firm. The contract status of a customer is pivotal in determining customer retention behaviour since, in a non-contractual

---

[15] Hazard rate represents the magnitude of risk that a customer will terminate their subscription at time (t). Further discussion is provided in Chapter 4, Section 4.3.

relationship, the entry dates are observed but exact exit dates are not observed (Fader, Hardie, & Lee, 2005). Consequently, customers in a non-contractual relationship are more susceptible to defection as they can do so without incurring switching fees.

3.5.2 *Explanatory Variables*

3.5.2.1 Customer transaction explanatory variables

Communication encounter:

The firm captured and classified customer-firm interactions as one of three categories: *inquiry, feedback* and *customer request*. Communication encounters are operationalised as time-varying variables. That is, the frequency of communication encounters for each customer is included throughout the duration of the study. Frequency is used as a measure for operationalising communication encounters because intensity of interactions is a basic dimension and key determinant of relationship continuity.

Length of relationship:

The length of relationship is the duration of a customer-firm relationship. Operationalization of the length of relationship is essential because of the lack of a formal contract. As noted earlier, customers in this category can move to another firm without switching fees. Length of relationship is used to understand the effect of satisfaction on relationship duration (Bolton, 1998). In this research, the variable 'length of relationship' was created for each customer and observed throughout the study, based on acquisition and departure dates. The value of these time varying variables are measured and updated monthly for the study duration, while a customer remains with the firm. The unit of analysis for this variable is "time in months".

Usage:

In relationship marketing literature, usage is measured in different ways. For example, Venkatesan and Kumar (2004) use the number of units purchased as the proxy for usage. Conversely, Nitzan and Libai (2011) associate usage with average monthly hours. Borle, Singh, and Jain (2008) adopted actual purchase amounts as their general measurement for modelling usage. Usage in this study is operationalised as a time-varying measure and because the nature of the products differs, number of "physical downloads" is used for Internet and "minutes" in the case of voice. The value of these time varying variables are measured and updated monthly for the duration of the study while a customer remains with the firm.

3.5.2.2 Customer complaints explanatory variables

Complainants versus non-complainants:

As proposed by Kau and Loh (2006), this research separately groups customers who have complained and those who have not. One reason for this separation is to illustrate the effect of complaining behaviour on customer retention. The variable is coded as "0" if the customer has not complained during the study and "1" if the customer has complained one or more times. For dichotomous or categorical variables one unit change of this variable is compared to the omitted reference category (Allison, 2010). In this study, the complaints are the reference group and the results will be compared to the non-complainants.

Number of complaints:

Most existing studies explore the complaint number variable through use of questionnaires gauging satisfaction or dissatisfaction. For example Sharma et al. (2010) found that greater levels of dissatisfaction led to greater numbers of complaints, regardless of the background culture. On the other hand, Ahn et al. (2006) determined dissatisfaction using quality of call drops as a way to understand customer churn. In this study, the complaint number variable reflects “the number of times each customer has complained during the study”. The variable ‘number of times’ measured in this study refers to the number of episodes of complaints emanating from the customers. The reason for such calls may be about technical related issues, billing or pro-longed delay in provisioning activities. The variable is recorded as a time-varying measure, is updated monthly, and varies among customers based on number of complaints.

Time since last complaint:

The variable of time since the last complaint measures the time between first and last complaint for each complaining customer. While customer A and customer B may make the same number of complaints, the time interval between these complaints may differ. Highly concentrated complaints are likely to be more influential in defection than complaints of lower concentration. The likelihood of defection for customer A and customer B are likely to vary because, combined with frequency of complaints, the “recency” of complaints may also be a factor. Following the conclusion by Bolton et al. (2006) that recent experiences weigh more heavily than early ones, this variable is recorded as a time-varying measure of “numbers of months between two consecutive complaints". This number is updated monthly throughout the customer-firm relationship during the observation period.

Severity of complaints:

Reasons for customer complaints include requesting refunds, reporting poor service quality, raising technical concerns, and addressing billing issues. However, not all service failures have the same severity, just as not all customers complain. A common method of measuring complaints is to categorise failures as either of high or low severity (McQuilken, 2010). Conversely, Conlon and Murray (1996) used product price and level of dissatisfaction as a surrogate to measure problem severity. In this research, direct measures from customer complaint datasets are used to operationalise the severity of complaints. For example, when a firm receives a complaint, severity is immediately assessed and determined to be low or high. While the severity of a complaint is subject to employee judgment, it is assumed that the firm's response will be influenced by its perception of the complaint's legitimacy.

Furthermore, the assumption is reasonable as Firm A has carried out customer satisfaction research since 2008 through internal sources and external research[16]. According to Firm A, their external research firm used a multi-modular approach to carry out customer satisfaction research including face-to-face interviews, computer assisted telephone

[16] Firm A have engaged TNS, a global research agency, to carry out research on a Customer Satisfaction Index (CSI) across all lines of their businesses which includes the consumer and business segments. As the materials and literature are copyrighted by the agency, information about the structured questionnaires used in the survey is not included in this thesis. Firm A confirms that the respondents to the survey were subscribers for at least for a period of six months or more.

interviews (CATI) and surveys. The collection of this data by the research firm provided the basis for establishing the quality dimensions and attributes for the Firm A's complaints measurement. Firm A emphasizes the importance of this measurement in their training programs for their customer contact centre. The aim of this training is to increase professionalism and methodological competencies while minimising errors during the service process. Thus, the assumption that the firm can make judgements about complaint severity and the customer contact centre will be fair in their evaluation of the severity of the complaints is considered reasonable.

The variables are recorded categorically as "high", "low", or "no complaint". In this study, customers with "no complaint" are the reference group and the results will be compared to complainants with "high" and "low" severity.

Length of recovery:

Timeliness of service recovery is a key measurement supported in the study by Davidow (2003). In most research, the recovery effort has been operationalised within the context of customer satisfaction. Customer satisfaction is determined by customer surveys that gauge whether a customer intends to continue their relationship with a firm (Kau & Loh, 2006; Smith et al., 1999). The variable to support the timeliness or speed of recovery was measured as a dichotomous response, namely, "0" (if delayed) and "1" (if immediate). However, this type of classification does not accurately measure recovery efforts, as most service firms have established protocol for managing recovery efforts and limited means. In this research, the "number of days taken to recover from a complaint" is captured as a time-varying measure, and is updated monthly for the period of customer-firm relationship during the observation period.

3.5.2.3 Customer characteristics

Age:

This variable is categorised as one of four age groups (i.e., 19-25 coded as “1”, 26-40 coded as “2”, 41-55 coded as “3”, and over 55 coded as “4”). The age group “1” is the reference group. Although age could be treated as a continuous variable, as categorical variable it allows for measurement of customer defection behaviour by lifecycle stages. The classification of age is similar to the existing research by Keaveney and Parthasarathy (2001); Mittal and Kamakura (2001).

Gender:

Gender is included as one of the key demographic variables in this study. The operationalisation of gender in this study is similar to that of Ahn et al. (2006) and Nitzan and Libai (2011). This variable is categorised as “F=Female” and “M=Male”. Female is the reference group. While there maybe differences in gender orientation, complaints and defection, this relationship is difficult to examine in this study. Lack of demographic information such as occupation, marital status, and educational level in the dataset precludes from accurately assessing the differential impact of these potential variables between complaining and defection.

Household income:

This variable is categorised as one of three income groups (i.e., “low”, “medium”, and “high”). The low income group is the reference group. The operationalisation of income is from the firm’s database and analogous to research by Reinartz and Kumar (2003). The income from the firm’s database has three different scales from low to high and follows the country’s census information.

Household size:

This variable is categorised as one of four household sizes (i.e., “1-“, “3-4”, “5-6”, and “more than 6”). The group “1-2” is the reference group. The operationalisation of household size is from the firm’s database and scales are coded using firm’s existing classification. The use of this scale is reasonable since the firm consistently carried out market research to evaluate the number of active household members. Therefore, the scale applied in this research is captured based those four levels.

The reference groups were chosen as the lowest numbers or first category to facilitate ease of interpretation. This is particularly helpful when dealing with complex models with interactions (Allison, 2010). A summary table of dependent and explanatory variables is presented in Table 3.2 below.

Table 3.2: Summary of Dependent and Explanatory Variables

Variables	Description	Unit of Analysis	Support
Dependent variable	The outcome variable measuring customer survival probability.	Not applicable	Allison, 2010; Cox, 1972)
Customer complaints information			
Complainants and non-complainants	The variable differentiate between customers who have logged complaints or otherwise.	0= non-complainants 1= complainants	Hadden et al. (2006); Kau and Loh (2006)
Number of complaints	The variable captures how many times the customer has logged a complaint with the firm over the duration of the relationship with the firm.	Number	Ahn et al. (2006)
Time since last complaints	The variable captures the effect "recency" of the previous complaint logged to the firm over the duration of the relationship with the firm.	Months	Bolton et al. (2006)
Severity	The variable captures severity of the complaint based on the service representative's judgment.	0= none 1= low 2=high	Smith and Bolton (1998)
Length of recovery	The variable captures the time taken by the firm to restore the failure.	Number of days	Knox and Van Oest (2014)

Variables	Description	Unit of Analysis	Support
Customer transaction information			
Length of relationship	The variable captures the number of months a customer has been with the firm.	Months	Nitzan and Libai (2011)
Usage	The variable captures physical download (for Internet) and number of minutes (voice) for each customer over the duration of relationship with the firm. It reflects the utility of the service to the customer.	Number of minutes/download	Ascarza and Hardie (2013)
Communication encounter	The variable captures any form of communication activities to connect with each of the customers throughout the journey from initial contact.	Number of contacts	Lemke et al. (2011)
Customer characteristic information			
Age	The variable captures age of the customers.	1=19-25 years old 2=26-40 years old 3=41-55 years old 4= over 55 years	Baumann et al. (2005)
Gender	The variable captures the gender of the customers.	1= Female 0= Otherwise	Melnyk et al. (2009)
Household Income	The variable captures the household income of the customers.	1= Lower 2=Medium 3=Higher	Reinartz and Kumar (2003)
Household Size	The variable captures the number of members living within a single household.	1= "1-2" 2="3-4" 3= "5-6", and 4="more than 6"	Mittal and Kamakura (2001)

3.6 Data Cleaning and Description

Examining the data values of the identified variables is an essential starting point in ascertaining the characteristic shape of the distribution. This includes checking the attributes of the categorical and continuous variables, and evaluating the extreme missing values and unexpected errors in the dataset (Berry & Linoff, 2004). For example, duplicate variables with text and address are not useful in understanding customer retention behaviour. In addition, a variable such as gender should have one of only two possible values (i.e., female or male), and any other values should be eliminated. Likewise, extreme missing values could lead to parameter estimate bias of the explanatory variables.

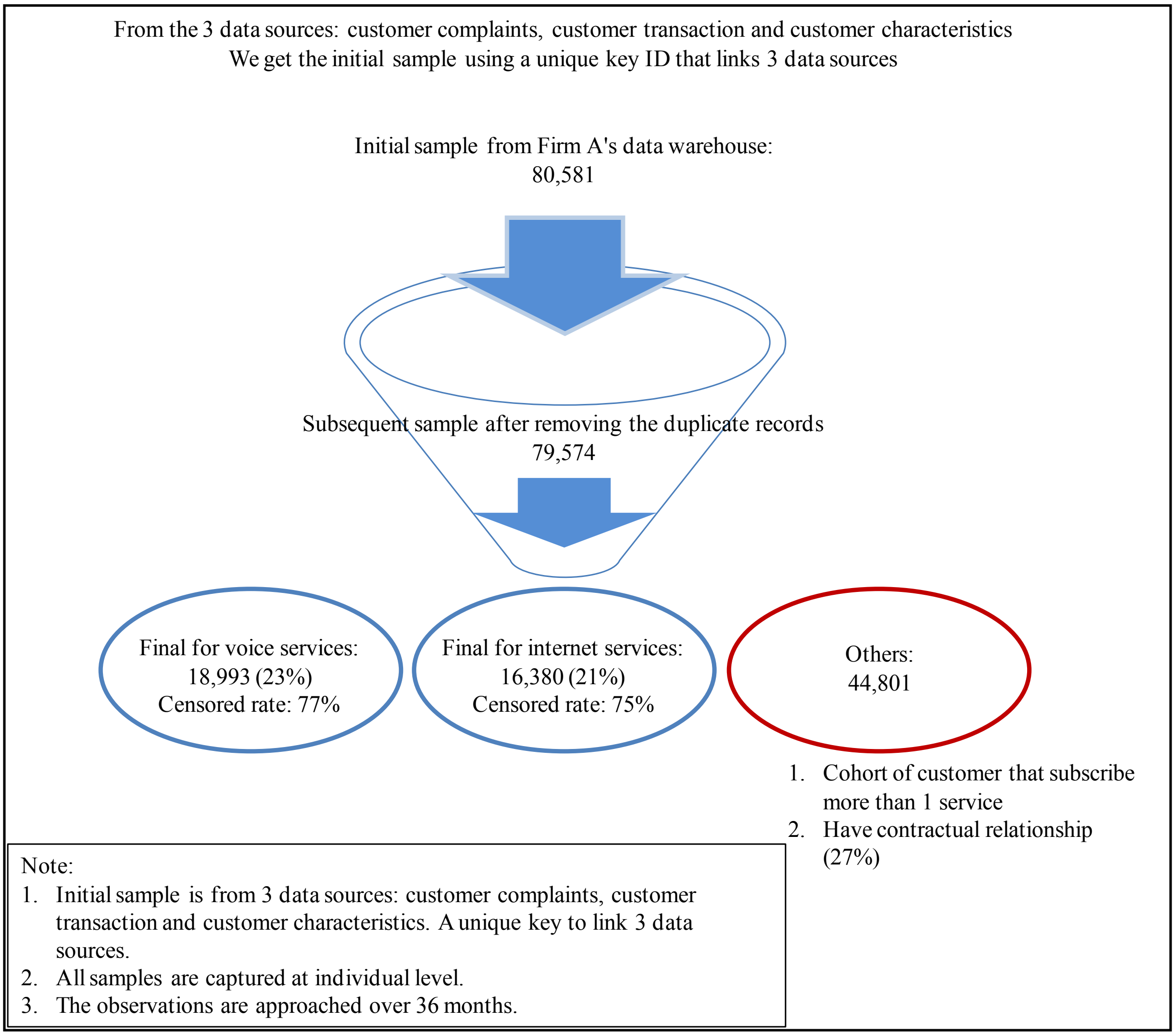
To prevent possible bias, missing observations were first examined in the three databases and variables with substantial missing data points were dropped. The first step in the data cleaning process is a case-wise deletion approach to remove variables for which more than 50 per cent of the data were missing. Although this approach is simple, it may seem wasteful. However, too many missing values results in inaccurate analysis. Significant absences may inflate the parameter estimates, increase standard error, and may lead to non-convergence (Van Buuren, Brand, Groothuis-Oudshoorn, & Rubin, 2006). Furthermore, when the percentage of missing values increases to more than 50 per cent, the overall model performance deteriorates (Kamakura, Wedel, De Rosa, & Mazzon, 2003). The second step of the data cleaning process is removing duplicate variables and identification variables not used in the model. In all three datasets, several identification and text variables, including address and product names, were made available. These types of variables are considered "noise" and are not useful in understanding customer retention.

Although the results of the model estimation are provided in the subsequent chapter, it is useful to examine the sample characteristics and descriptive statistics for the explanatory variables included in the model at this point in the thesis.

The initial sample characteristic contains 80,581 individual customers extracted from the firm's data warehouse. After removing duplicate records, total customers who were active on 1st January, 2010 are 79,574 customers. Because this research examines non-contractual customers only, further reduction was necessary. Among these customers, 23 per cent and 21 per cent of customers are subscribers to only voice or Internet services respectively. The remaining 56 per cent of customers either subscribe to more than one service or are in a contractual relationship. This narrowing is imperative because multi-service customers are less likely to defect (Kamakura et al., 2003).

Additionally, even if the customers terminate one service, their defection is most likely considered switching behaviour, rather than simply leaving the firm and discontinuing the use of the service overall. Similarly, non-contractual relationships must be isolated because these customers can leave the firm without incurring an exit fee. Last, the censoring information for customers subscribing to voice and Internet services is 77 per cent and 75 per cent, respectively. This means these customers have not defected upon completion of data collection, so we do not know the true duration of the customer-firm relationship. The data sources used in this research are shown in Figure 3.1 below.

Figure 3.1: Data sources in this study



Overall, there are twelve explanatory variables examined within this study. These explanatory variables, which consisting of categorical and continuous data types, are included in the model estimation. The overall means and/or frequencies are provided shortly in Tables 3.3, 3.4 and 3.5.

3.7 Descriptive Results

The descriptive statistics for the explanatory variables included in the model are presented in Tables 3.3, 3.4 and 3.5[17]. The overall means and/or frequencies are listed as appropriate for customers subscribing to voice and Internet services[18]. It is important to consider the characteristics of these explanatory variables at an early stage. The sample sizes for voice and Internet subscribers of 18,393 and 16,380, respectively, are large enough to estimate customer retention behaviour (Blattberg et al., 2008; Neslin, Gupta, Kamakura, Lu, & Mason, 2006).

As highlighted in earlier chapters, the model is estimated using both categorical and continuous explanatory variables. Table 3.3 presents information relating to categorical explanatory variables. Tables 3.4 and 3.5 relate to continuous explanatory variables for customers subscribing to voice and Internet services.

For voice services, Table 3.3 shows the number of complainants versus non-complainants to represent 30 and 70 per cent of customers respectively. For Internet services, Table 3.3 shows the number of complainants versus non-complainants to represent 21 and 79 per cent respectively.

As pointed out in section 3.5.2, customers who have complained one or more times, are flagged as complainants. Further analysis from the complainants provides the information related to complaint severity. This analyses shows that customers who experienced low and high severity incidents for voice services are 3 and 5 per cent respectively. For

[17] The graphical examination of the continuous explanatory variables is presented in Appendix A.

[18] The correlation matrix for the continuous explanatory variables can be found in Appendix B.

Internet services, low and high severity incidents are 2 and 6 per cent respectively. The analysis was carried out using the frequency procedure in SAS[19].

[19] The Frequency procedure produces frequency and cross tabulation table that helps to measure the association across, as well as within, the levels (Allison, 2010).

Table 3.3 : Descriptive Statistics for Categorical Variables

	Product: Voice		Product: Internet	
Explanatory variable	**Frequency**	**Percent**	**Frequency**	**Percent**
Age				
Group 1 (19-25)	1,213	7%	1,715	10%
Group 2 (26-40)	5,616	31%	5,510	34%
Group 3 (41-55)	7,590	41%	6,614	40%
Group 4 (over 55)	3,974	22%	2,540	16%
Gender				
Male	11,582	63%	9,217	56%
Female	6,724	37%	7,111	44%
Household income				
Low	1,680	9%	1,572	10%
Medium	13,396	73%	11,816	72%
High	3,317	18%	2,990	18%
Household size				
01-02	4,263	23%	3,857	24%
03-04	8,034	44%	6,987	43%
05-06	4,881	27%	4,450	27%
Over 6	1,215	7%	1,086	7%
Complainants	198,432	30%	126,033	21%
Non-complainants	463,716	70%	463,677	79%
Severity				
None	618,164	93%	547,505	93%
Low	10,145	2%	11,312	2%
High	33,839	5%	30,893	5%
Note: Sample size for voice = 18,393 customers Sample size for internet = 16,380 customers				

Four customer characteristic variables were included in the model. First is gender. The model shows voice customers are 63 per cent male and 37 per cent female. Internet customers are 56 per cent male and 44 per cent female.

Age is segmented into four categories. Group 1, customers between 19 and 25 years of age, represents 7 per cent of voice and 10 per cent of Internet subscribers. Group 2, customers between 26 and 40 years of age, represents 31 per cent of voice and 34 per cent of Internet subscribers. Group 3, customers between 41 and 55 years of age, represents 41 per cent of voice and 40 per cent of Internet subscribers. Finally, group 4, customers over 55 years of age, represents 22 per cent of voice and 16 per cent of Internet subscribers.

Household income is segmented into the three categories of low, medium, and high. Around 9 and 10 per cent of customers are in the low income categories of voice and Internet services respectively. 73 and 72 per cent of customers are in the medium income categories of voice and Internet services respectively. For, both voice and Internet services, high income customers represent 18 per cent of subscribers.

Finally, household size is segmented into four categories. Customers with 2 or fewer members of a household account for 23 and 24 per cent of voice and Internet subscribers respectively. Customers with 3 to 4 members of a household represent 44 and 43 per cent of voice and Internet subscribers respectively. Both categories of 5-6 and more than 6 member households represent 27 and 7 per cent of voice and Internet subscribers respectively.

Tables 3.4 and 3.5 below show the descriptive statistics of continuous variables for voice services and for Internet services.

Table 3.4: Descriptive Statistics for Continuous Variables: Voice Services

Explanatory variables	N	Mean	Std Dev	Minimum	Maximum
Number_of_complaints	18,393	0.050	0.333	0	24
Usage	18,393	101.345	112.706	0.01667	500
Time_since_last_complaint	18,393	11.155	9.566	0	35
Length_of_relationship	18,393	175.410	95.777	2.66667	477
Communication_encounter	18,393	0.156	0.804	0	48
Length_of_recovery	18,393	0.175	2.272	0	576

Table 3.5: Descriptive Statistics for Continuous Variables: Internet Services

Explanatory variables	N	Mean	Std Dev	Minimum	Maximum
Number_of_complaints	16,380	0.053	0.372	0	30
Usage	16,380	25.469	44.442	0	499.36
Time_since_last_complaint	16,380	10.220	9.067	0	35
Length_of_relationship	16,380	161.756	96.673	3	477
Communication_encounter	16,380	0.166	0.878	0	90
Length_of_recovery	16,380	0.190	2.497	0	1152

Tables 3.4 and 3.5 above illustrate standard deviation for all continuous variables in product categories, voice and Internet, as larger than the mean. The spread of data across all variables may account for this deviation. Averaging across the time period and customers of the dataset may also play a role (Reinartz & Kumar, 2003).

3.8 Summary

This chapter provides an overview of the context of the research. The uniqueness of the data stems from the three sources of databases from Firm A. The willingness of Firm A to provide data for two products at different product life stages provides a basis for modelling customer retention behaviour overtime in two contrasting market contexts. Discussion with Firm A's representatives provided further insights relating to the explanatory variables.

The next chapter describes technical aspects of the methodology used in this study. It describes several technical methods that can be used to model customer retention and explains why a survival analysis methodology is the most appropriate method to be utilised in this that study. Three forms of survival analysis that permit analysis of a set of explanatory variables are outlined, including: non-parametric, parametric and semi-parametric approaches. The choice of one form of semi-parametric approach, the Cox Proportional Hazard Model, is justified as this method best supports the study's research aims and the structure of the research data. Estimation procedure and the underlying mathematical specification are also described.

CHAPTER 4: METHODOLOGY

4.0 Introduction

The aim of this study is to understand how the effects of various explanatory variables influence customer retention behaviour. The research compares the survival probability of customers at individual levels for two products, one at the decline stage in the product life cycle and one a growth stage of the product life cycle. Further, customer retention behaviour is estimated to capture unobserved heterogeneity using random effects.

The methodology used to address this research aim is described in detail in this chapter. First the chapter discusses existing modelling techniques and approaches to estimating customer retention behaviour including: decision trees; the "recency", frequency and monetary (RFM) model; hazard models; regression; and logistic and survival analysis. Next, this chapter considers the appropriateness of a survival analysis approach to this research and why the Cox Proportional semi-parametric model is best suited to and adopted for this study. The model equation and estimation procedure adopted for the research are then discussed. A discussion of the model application then shows how explanatory variables are presented for three models: the base line, comprehensive and interaction model. This chapter then provides the model equations.

4.1 Modelling Techniques

An increasing number of firms are moving to a subscription-based service model with the aim of strengthening their revenue bases and offerings. As a result of this shift, much of their marketing strategy and efforts need to move from product to customer centric initiatives. This shift points towards a trend wherein marketing efforts need to focus on building the customer base, and a focus on customer retention that seeks to reduce defection rates. Essentially, the modelling approach to be adopted for a study of customer retention is contingent on the type of data, the nature of the research questions investigated and the significance given to the interpretation and/or its explanation against prediction. Models that may provide excellent explanations may not provide suitable predictions, primarily because of issues regarding data interpretation.

The two suggested approaches to help explain customer retention behaviour are the ‘single future period’ approach and the ‘time series’ approach (Blattberg et al., 2008). ‘Single future period’ is favoured where the research interest is customer behaviour at a specific time (Rindfleisch, Malter, Ganesan, & Moorman, 2008). Using cross-sectional data, this approach includes logistic regression, decision trees, discriminant factor analysis, and neural networks (Ahn et al., 2006; Homburg & Giering, 2001; Lemmens & Croux, 2006; Neslin et al., 2006).

Use of the ‘time series’ approach is best suited to understanding customer retention behaviour over a number of time periods (Van den Poel & Larivière, 2004). Focusing on longitudinal data, this approach includes customer life time value analysis, time series regression, hazard models and proportional hazard models (Bolton, 1998; Buckinx & Van den Poel, 2005; Danaher, 2002; Jamal & Bucklin, 2006).

Since this study seeks to understand customer retention behaviour over time, a ‘time series’ technique is appropriate. There are a number of ‘time series’ approaches used to examine customer retention behaviour and these are now reviewed.

4.1.1 *Decision Trees and RFM Analysis*

Classified as non-parametric methods, decision trees and RFM models are often used to understand customer retention behaviour at a single future period. In the context of understanding customer retention behaviour, a decision tree uses a binary tree-like model with branches that provide an easy identification of significant explanatory variables, and rules to predict the relationship outcome. Decision trees are an easily implementable approach requiring limited statistical assumption and can incorporate both categorical and continuous type explanatory variables. However, there are problems of alignment of this technique to the objective of this research. There are two key aspects often neglected when using a decision tree (Neslin et al., 2006). First, there is a loss of information because non-parametric methods rarely provide insight into the impact of each explanatory variable on the outcome variable. Second, an outcome variable can only have discrete values and this does not align with the continuous time outcome variables of this research. Given these limitations, decision trees do not align well with the objectives of this research's proposed model.

The RFM model analyses customers in terms of how recently they have purchased a product or service, how often they purchase the product or service, and how much money they spend on the product or service. Although this method is easy to implement, there are two problems with the RFM model. First, it limits the number of explanatory variables. Second, it assumes a homogeneous customer, which is contrary to the argument, in this research, of heterogeneous behaviour across customers. Also, with a major focus on past data, RFM does not show what is driving customers to make purchases. Therefore, the disadvantages of the RFM approach make this approach unsuitable for this research.

4.1.2 *Regression and Logistic Regression Models*

Both regression and logistic regression models are used to examine customer retention behaviour based on the inclusion of one or more explanatory variables. The regression model provides a simple way to examine if association exists between the explanatory variables and the relationship outcome (retention or defection). The logistic regression model examines the relationship between one or more explanatory variables and dependent variables, and presents the significance of their relationships. More specifically, in a linear regression model the dependent variable is considered continuous, while in logistic regression it is categorical or discrete. However, these conventional methods have difficulty in handling the two common features of the longitudinal dataset used for this research: censoring and including time varying explanatory variables (Helsen & Schmittlein, 1993).

As a result of the limitations of the models described above, the modelling approach determined as most appropriate for this study is survival analysis, which is now discussed.

4.1.3 *Survival Modelling and its Relevance in this Research*

The aim of this research is to understand the factors that influence customer retention behaviour by estimating customer survival probability. There are four important aspects with respect to the research design that need to be considered.

First, the research is designed to utilise a firm's longitudinal data to track a customer's retention behavioural pattern at multiple points in time over a specified period. With this, the temporal effect is taken into account.

Second, the research design involves use of explanatory variables derived from the three databases within the firm: customer complaints, customer transactions, and customer characteristics. The research has focused on modelling a range of explanatory variables to demonstrate the differences of customer retention behaviour over a period of time. For instance, one of the key drivers of customer retention could be the duration of the relationship between the firm and customer's survival probability. However, other explanatory variables, such as usage or number of complaints, could also explain customer survival probability at different points of time.

Third, a further aspect of this design is that a customer relationship with the firm starts and ends at different points during the observation period. In other words, a customer's behavioural pattern and survival time are not uniform across other observations - some customers may exceed the observation time. In this study, this research observed the relationship between customers and the firm from January 1st 2010 to 31st December 2012 inclusive. For each customer relationship there are two possibilities:

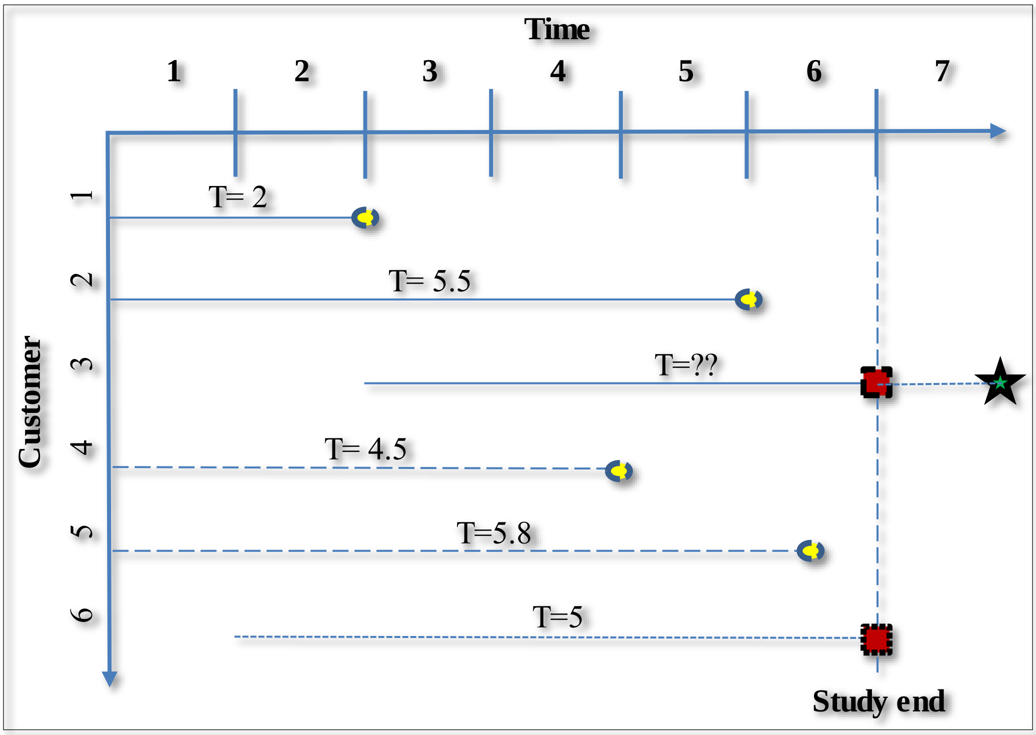
(a) The relationship between the customer and firm terminates between January 1st 2010 and 31st December 2012. Customers who experienced this event[20] are called the observed cases and we have their observed duration; and

---

[20] Event could be customer defection or customer making subsequent purchase.

(b) The relationship between the customer and firm still exists after 31st December 2012, beyond the research observation period. Here, the study can only conclude that the relationship has not ended and the true duration of their observation remains unknown. It is not known if the customer survives after the observation period. This is called censored data and it is 'right censored' if the end of the relationship time is unknown. Conversely, if it is the start of the relationship that is not observed it is called 'left censoring' (Allison, 2010). Figure 4.1 below further demonstrates the censoring concept.

Figure 4.1: Censoring



Source: Adapted from Allison (2010)

In the above Figure, there are six customers observed over time from 0 to 6. The solid line represents an observed period at risk while the small circle represents an observed event. An observed event in this research is customer defection or survival at the particular time T. Survival time is recorded for each customer. The broken line represents an unobserved period at risk while the square box represents the censoring time. The star represents an unobserved event. For customer 3, the observation 3 is censored in that it has not experienced the event at the end of the observation period, so its true duration is unknown.

Censoring occurs when the information about the customer's survival time is incomplete. One of the underlying reasons to include censored data in the research is to avoid potential bias arising from spurious duration dependence (Van den Bulte & Iyengar, 2011). This is especially the case when the number of customers with complete observations is smaller compared to incomplete observations, as it represents a particular type of missing data. Ignoring this would impede the calculation of the parameter estimates. Therefore, an appropriate method is required to effectively handle both censored and uncensored observations.

Fourth, and finally, the research data contains time invariant and varying explanatory variables. Time invariant explanatory variables are constant inputs throughout the research such as gender or race. Time varying explanatory variables include the number of complaints, and age, both of which may have changed through the duration of this research. For example, the number of complaints for each of the customers is different and they occur at different points of the duration of observation. Each explanatory variable has a value per customer as well as multiple values per customer. Therefore this research requires a model that is able to estimate both *time invariant and time varying explanatory variables*, creating uniqueness to the research model requirements.

More generally, given the uniqueness of the marketing data in this research, there is a need to apply a model that is able to decipher time invariant and variant aspects of a customer retention pattern. Survival models represent a suitable approach to analyse this type of duration data (Helsen & Schmittlein, 1993) and the next sub-section will argue the appropriateness of survival analysis for use in this research.

4.2 Survival Analysis

Survival analysis examines the data of the 'outcome variable' until the occurrence of an event of interest (Allison, 2010; Helsen & Schmittlein, 1993). The outcome variable in this research is the probability that the relationship between customer-firm survives in a series of explanatory variables. The time for the event, or survival time, for each of the customers can be measured in days, weeks, months or years. This research will focus on 'months' as the time for the event, as this aligns with the firm's monthly data collection procedures.

Unlike conventional modelling techniques, survival analysis is designed to support temporal variation in the data used in this research. As noted earlier, observations can have different time variants during the period of this research. As such this method is advantageous when compared to approaches such as linear regression or logistic regression (Lawless, 2011).

As the aim of this research is to understand factors that impact customer retention behaviour, the research design includes a range of explanatory variables using different databases.

A crucial assumption of this research is that customer retention behaviour can no longer be predicted by using only demographic variables alone, such as age, gender, marital status. Demographic variable models oversimplify the complex customer - firm relationship (Jamal & Bucklin, 2006). This research is designed to analyse customer survival probability being contingent upon several factors with explanatory variables such as customer touch points to provide additional insights into customer retention behaviour (Braun & Schweidel, 2011). The parameters of focus of this research are time invariant and variant data. Unlike conventional methods, survival analysis incorporates time varying explanatory variables as well as censoring information from both censored and uncensored observations. Table 4.1 summarises the comparisons drawn of the relevant marketing models and the overall suitability decision survival analysis.

Table 4.1: Suitability of Relevant Models

No.	Modelling Technique	Explanatory Variables	Outcome Variable	Censoring Permitted
1.	Linear regression	Categorical or continuous	Normally distributed	No
2.	Logistic regression	Categorical or continuous	Binary or ordinal or nominal	No
3.	Decision Tree/RFM	Categorical or continuous	Binary or ordinal or nominal	No
4.	Survival analysis	Time categorical or continuous	Binary or ordinal or nominal	Yes

Source: Summarised from Blattberg et al. (2008)

4.2.1 *Building Blocks of Survival Analysis*

Having established survival analysis as the appropriate approach, this sub-section will describe the building blocks of survival models which will be used in this research.

As discussed above the dependent variables in survival analysis comprise two parts: one is the time to event; and the other is the event status, which records if the event of interest has occurred or not. Based on this, one can estimate two functions that are dependent on time, the survival and hazard functions.

The survival and hazard functions are key concepts in survival analysis for describing the distribution of event times. The survival function, denoted as $S(t)$ gives, for every time, the probability of surviving (or not experiencing the event) up to that time. The hazard function, denoted as $h(t)$ gives the potential that the event will occur, per time unit, given that an individual has survived up to the specified time. While these are often of direct interest, many other quantities of interest (e.g., median survival) may subsequently be estimated from knowing either the hazard or survival function. It is quite normal for survival studies to describe the relationship of a factor of interest (e.g., repurchase intentions) to the time to event, in the presence of several covariates, such as age, gender, race, time etc. (Helsen & Schmittlein, 1993).

For example, within this research one can consider the "survival model" as the probability that a firm retains a customer (or loses one) as a function of time. In this situation, $t = 0$ is the moment in which the firm acquires the new customer and starts with the probability of $p = 1$ of retaining the customer. As t increases in time ($t > 0$), then the probability of retaining a customer decreases.

4.2.2 *Types of Survival Analysis Approaches*

There are three types of approaches in survival analysis that enable analysis of the relationship of a set of explanatory variables: they are non-parametric, parametric and semi-parametric approaches. The selection of why one approach is better than the others is conditional on the research questions and type of data structure. As highlighted earlier, the aim of this research is to understand customer retention behaviour over time using not only customer transaction and customer characteristics but specifically customer complaints. Given the relative importance in this research in understanding the factors that explain customer retention behaviour, the decision to select the modelling approach will be significantly dependent on the importance between the explanation and prediction, due to the fact that models that perform well at explanation may not necessarily do well at prediction (Kamakura et al., 2005).

In addition, the type of data structure is pivotal in selecting appropriate survival approaches. To recapitulate, this research also aims to estimate survival probability on individual-level data. Because of the unique facet of the three databases used for this research, the models must be estimated on person-period data (Singer & Willett, 2003). The data in this study is organised with one row of data per customer, with one dependent variable representing the time to event; and another one to indicate if event occurred or censored.

The next section provides further discussion on each of the non-parametric, parametric and semi-parametric approaches to survival analysis.

4.2.2.1 Non-parametric

A non-parametric approach can be used to model customer-firm relationships without assuming any baseline hazard distribution, or how the explanatory variables affect the shape of the hazard function (Allison, 2010). In other words, one can estimate the survival function or survival probabilities of a customer as a function of time. This type of approach is best suited to understand 'natural patterns' in customer survival, and could be used to identify key points where survival rates fall. One of the well-established approaches is the Kaplan Meier curves and life-table methods (Gilula, McCulloch, & Rossi, 2006).

Whilst relatively easy to estimate and interpret, non-parametric approaches do not allow for the inclusion of multiple explanatory variables and multivariate controls, required for this research to understand customer retention behaviour. Omitting them will result in loss of information since variance between the parameters cannot be incorporated. Another problem with non-parametric approaches is the inability to detect shifts in the mean or the relationship between several related explanatory variables. For these reasons a non-parametric approach is determined unsuitable and the discussion now moves to the parametric approach.

4.2.2.2 Parametric Approach

Under parametric approaches, one must decide in advance the shape of the underlying distribution of the hazard function. Accordingly, how the explanatory variables impact the baseline hazard is assumed to vary in a specific manner and follow certain known probability distributions (Singer & Willett, 2003). For example, 'Weibull' assumes a monotonic distribution (decrease or increase), ‘Exponential’ assumes constant distribution, ‘Gompertz’ assumes an exponential increase, and ‘Log-Logistic’ assumes a non-monotonic hazard.

The parametric approach allows easy incorporation of explanatory variables and calculates rates, smooth and nuisance data. Although this approach provides precise parametric estimates, and allows for multivariate analysis, it imposes the most structure of those variables that are considered. As a result, the researcher needs to be sure that the estimated distribution matches the data structure. In many real world cases the appropriate distribution is not known (nor are they able to be closely approximated). With censoring and the explanatory variables we cannot do a simple histogram and say "This looks like a .... distribution to me".

Since this current study deploys actual customer behavioural data, it may be sensitive to the inclusion and omission of the explanatory variables. Furthermore, incorrectly specifying the hazard function may lead to incorrect parameter estimates and can result in serious bias. Another limitation of the parametric approach is that it does not allow time-variant variables (Nitzan & Libai, 2011). Considering the unique marketing data and the temporal effects in this dataset, the parametric approach is not deemed suited to this current study.

4.2.2.3 Semi-parametric Approach

A semi-parametric model makes fewer assumptions than typical parametric methods, but more assumptions than the non-parametric methods described above (Allison, 2010). The advantage of semi-parametric approaches is their ability to accommodate survival models without knowledge (or assumptions) of the hazard distribution. More importantly, the semi-parametric model has the ability to incorporate time-varying explanatory variables. As noted earlier, the data for this research is unique, involving multiple records of customers with multiple explanatory variables over 36 months. The longitudinal-data situation requires flexible semi-parametric approaches and this presents an overall more robust technique, compared to the parametric and non-parametric approaches. One particular type of semi-parametric approach, the Cox Proportional Hazard Model, will best support this study's research aims and the structure of the research data. This model is now discussed.

4.2.2.4 Cox Proportional Hazard Model

The Cox Proportional Hazard (PH) model is described as a semi-parametric model and makes no assumption about the shape of the distribution (Cox & Oakes, 1984). However, with the Cox PH model, the assumption of the model must be tested in advance. This is called the proportional hazards assumption. Ascertaining whether this assumption is met is important because this model makes no assumption about the form of hazard function, which is denoted as $h(t)$.

Detailed discussion of this assumption is provided in Section 4.5. This is considered the non-parametric component of the model. However, it assumes parametric form for the effect of the explanatory variables on the hazard. Considering the aim of the research is to understand the relative importance of the explanatory variables on customer retention behaviour, the study is more focused on parameter estimates than the shape of the hazard.

Further, the Cox PH model can handle censored data and stratification. In stratification, a categorical variable is used to create separate baseline hazards. The main advantage to this stratification is that the time invariant explanatory variables are not subject to the proportional hazards assumption, and that the stratification feature of this model allows for division and sub-division between different groups, thus lending itself to more in-depth analysis (Spicer, 2005). The flexibility of the Cox PH model to manage large numbers and types of explanatory variables explain the reason for the adoption of this approach in this thesis.

Finally, to capture the unobserved heterogeneity, the Cox's PH model will be extended by including a cluster-specific random effect, in the expression of the hazard function. Correspondingly, the random effects are used to model the association between individual survival times within clusters which allows correlated observations into proportional hazards models (Allison, 2010).

The subsequent sub-sections will describe the model equations, the estimation procedure, the examining of proportional assumptions for the Cox PH model and the model application. First, the model equations explain the underlying concepts of survival analysis which leads to specific equations in the Cox PH model. Second, the estimation procedure of the proposed model is delineated. Third, relating to testing the proportionality assumption which is central in running the Cox PH model is discussed. Finally, the model application builds upon the identified explanatory variables discussed in Chapter 3, Sub-section 3.5, which are utilized in the Cox PH model in this study.

The elements covered in the model provide a framework for evaluation and interpretation of the results, which aim at answering the eight research questions identified in Chapter 1.

4.3 Model Equation

There are several ways to characterize the probability distribution of a survival random variable which are specific to survival analysis. The following key terms adapted from Allison (2010); Cox and Oakes (1984); and Helsen and Schmittlein (1993) are summarised as follows:

Probability density function: T is a continuous positive random variable denoting the survival times ($T = Y$). For example the time customer will remain with the firm. T has a probability density function $f(t)$: roughly probability that an event occurs at time t. Because $F(t)$ is continuous, to get the probability, we would have to find the area under the density for an infinitely small interval around

Survival function: is the probability that a customer will have a survival time greater than or equal to t and denoted as $S(t)$. Since $f(t)$ has a corresponding cumulative density function:

$$F(t) = \int f(u)du = \Pr(T \leq t) \tag{1}$$

But in survival analysis the focus will be on the survival function denoted as $S(t)$, which is the probability of an event occurring by time t. Then the survivor function is the probability of surviving until at least time t expressed as:

$$S(t) = 1 - F(t) \tag{2}$$

Hazard function: is the risk of an event in a time interval after time t, given that the customer has survived to time t denoted as $h(t)$. For example, in the context of this research it helps to understand that given a customer has survived 1 year, what is the probability they will defect.

It is therefore expressed as:

$$h(t) = P(event\ at\ t|survival\ up\ to\ t) \tag{3}$$

$$= \frac{P(survival\ up\ to\ t|event\ at\ t)P(event\ at\ t)}{P(survival\ up\ to\ t)}$$

$$= \frac{P(event\ at\ t)}{P(survival\ up\ to\ t)}$$

$$= \frac{f(t)}{S(t)}$$

$$f(t) = h(t)S(t)$$

The survival & hazard function is convertible, that is, if you are using the survival function, you would get the probability of customer surviving beyond the observation time. From the hazard function, one could deduce the hazard that customers defect at time t. For parametric approaches, the hazard rate is considered as function of time; but, since this research is using the semi-parametric model, no assumption about the form of $h(t)$ is required. Thus, the model equation that forms the Cox PH[21] model incorporating the explanatory variables is expressed as follows:

$$h_i(t) = h_0(t)\exp\left(\beta_1 x_{i_1} + \cdots \beta_k x_{i_k}\right) \tag{4}$$

Where:

$h_i(t)$ is the dependent variable (operationalised as the hazard rate at time t for customer i),

[21] The model equation draws heavily from Allison (2010)

x_1 *to* x_k are k explanatory variables or covariates, and

β_1 to β_k are the parameter estimates;

$h_0(t)$ is a baseline hazard function and is left unspecified. The baseline hazard function can be thought of as the hazard function for an individual whose explanatory variables all have values of 0.

Variables with positive parameter estimates (β values) are associated with increased hazard and decreased survival times, i.e. as the explanatory variables increases the hazard of the event increases and the predicted survival duration decreases. Negative parameter estimates (β values) indicate decreased hazard and increased survival times (Giudicati, Riccaboni, & Romiti, 2013).

As noted earlier, one key aspect of the Cox PH model is the inclusion of the time varying explanatory variables which can be expressed as:

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i_1} + \beta_2 x_{i_2}(t) \qquad (5)$$

Where:

$x_{i_2}(t)$ is explanatory variables that varies in value for each individual with time t.

x_{i_1}, is value of the time-fixed explanatory variables and the hazard at time t depends on value of x_1 at time t or

x_2 *(t)* is the value of the time-varying explanatory variables can be defined using information about the customers in the study prior to time t thereby allowing for lagged or cumulative values of some variables.

The equation then can be expressed in the following form by taking the logarithm of both sides of the equation.

$$\log h_i(t) = \alpha(t) + \beta_1 x_{i_1} + \cdots \beta_k x_{i_k} \tag{6}$$

Whereas, log $h_0(t)$ is termed as baseline hazard but using Cox proportional hazard model estimates of the parameter (β) coefficients could be handled without having to specify the hazard rate function:

$$\frac{h_i(t)}{h_j(t)} = \frac{\cancel{h_0} \exp(\beta_1 x_{i_1} + \cdots + \beta_k x_{i_k})}{\cancel{h_0} \exp\left(\beta_1 x_{j_1} + \cdots + \beta_k x_{j_k}\right)}$$

$$= \exp\left[\beta_1 \left(x_{i_1} - x_{j_1}\right) + \cdots + \beta_k \left(x_{i_k} - x_{j_k}\right)\right] \tag{7}$$

In term 7, with this assumption, $h_0(t)$, the baseline hazard function cancels out from the formula expressing a hazard ratio for any two individuals *i* and *j*. Because $h_0(t)$ cancels out, a method termed as partial likelihood estimation, which discards the baseline function and treats only the second part of the equation is developed (Cox, 1972). This estimation will be described in the later section.

Finally, unobserved heterogeneity is another key aspect that requires attention in managing this type of data (Mills, 2011; Vaupel, Manton, & Stallard, 1979). For example, suppose we have j observations and i subgroups (for repeated measures data, the j observations will simply be the period-specific records of data for the individual). The hazard rate for the jth individual in the ith subgroup (with random effect) is expressed as:

$$h\left(t_{ij}\right) = h_0\left(t\right)\exp\left(\underline{\beta}^T \underline{x}_{ij} + \underline{\psi}^T w_i\right) \tag{8}$$

Where

w_i are the cluster of individuals,

$v_i = \exp\left(\underline{\psi}^T w_i\right)$ are the shared cluster of individuals.

Main difference between "shared" and "unshared" cluster is the assumption of how the random effect is "distributed" in the data. Shared- or group-random models assumes that similar observations share the same risk, even as that risk may vary from group-to-group (Jenkins, 1995).

4.4 Estimation Procedure

The basic idea under the Cox PH model is that information about parameter (β) can be obtained from the relative orderings (i.e., ranks) of the survival times, rather than the actual values. To explain the estimation procedure, recall equation (4) and the hazard function for the first subject can be expressed as:

$$h_i(t) = h_0(t)\exp\left(\beta_1 x_{i_1} + \cdots \beta_k x_{i_k}\right) \tag{9}$$

Where

x_1 represents the hazard value, x for customer 1.

Similarly, the hazard function for all the customers can be generated and expressed as:

$$h_1(t) + h_2(t) + \cdots h_n(t) \tag{10}$$

Therefore, the likelihood for customer 1 to experience event at time *t* is the ratio of hazard over the risk set (Cox, 1972).

That is:

$$L_1 = \frac{h_1(t)}{h_1(t) + h_2(t) + \cdots h_n(t)}$$

$$= \frac{\cancel{h_0} \exp(\beta x_1)}{\cancel{h_0} \exp(\beta x_1) + \cancel{h_0} \exp(\beta x_2) + \cdots + \cancel{h_0} \exp(\beta x_n)}$$

$$= \frac{\exp(\beta x_1)}{\exp(\beta x_1) + \exp(\beta x_2) + \cdots + \exp(\beta x_n)} \qquad (11)$$

The equation above cancels the baseline hazard and thus the likelihood function becomes as βx – the coefficient to be estimated with the explanatory variables. Remarkably, an important feature of the Cox PH model is even though the baseline hazard is unspecified, the parameter (β) can still be estimated while incorporating the censored information into the likelihood function. Although the resulting estimates are not as efficient as maximum-likelihood estimates for a correctly specified parametric hazard regression model, not having to make arbitrary and possibly incorrect assumptions about the form of the baseline hazard is a compensating virtue of Cox's specification

The partial likelihood technique is valid when there are no ties in the data set. Specifically, no two individuals in this research would have the same event time. This may not be realistic since in reality two or more individuals may have exactly the same value on event time. Hence, an appropriate method to handle the type of ties in data is needed. There are several ways to adjust the likelihood to take into account observed ties using either the Breslow or Efron approximations (Helsen & Schmittlein, 1993). This research will follow the Efron approximation as it provides better parameter estimates than the Berslow (Hertz-Picciotto & Rockhill, 1997). Furthermore, Breslow ties do not do well when the number of ties at a particular time point is a large proportion of the number of event at risk.

4.5 Examining the Assumption of Proportional

One of the important issues to be assessed before the model results can be applied is whether the proportional hazards assumption is plausible for the given data set. The key assumption in running the Cox PH model is that a proportional hazard exists. That is, the effect of a given explanatory variable is constant across time (Allison, 2010). The assumption is important since the Cox PH model relies on the premise that hazard function is constant over time. If the assumption is violated then the model is invalid and other modelling approaches would be required.

There are several ways to test for examining proportional assumption made under the Cox PH model[22]. First, by analysing the Kaplan-Meier Curves, which graph the survival function for two subgroups against survival time. If the explanatory variables satisfy the proportional hazard assumption, the shapes of the curves should be basically the same, and the separation between the curves should remain proportional across the duration of the analysis. However, this method does not work well for continuous explanatory variables or categorical variables that have many levels because the graph becomes too

[22] This information regarding the proportionality assumptions draws heavily on Allison (2000).

cluttered.

A second approach is to test the interaction effects of time varying explanatory variables as a function of time. If the test shows that interaction significantly exists, then the proportional assumption is violated.

Finally, the third and most popular assessment of proportionality is based on the Schoenfeld residuals, which should show no association with time, if proportionality holds. This requires testing time varying explanatory variables for a non-zero slope of the scaled Schoenfeld residuals as functions of time. If the residual exhibits an unsystematic pattern at each failure time, then this gives evidence that the explanatory variables are not changing with respect to time—hence the PH assumption holds. Conversely, if the residual exhibit systematic pattern, it indicates that as time passes, the effects of the explanatory variable is changing. Therefore, the PH assumption does not hold because it displays temporal trends.

In summary, there is a constant non-negative acceleration factor that stretches out or shrinks survival times (Blattberg et al., 2008; Buckinx & Van den Poel, 2005). If the proportional hazard condition is satisfied, then it becomes possible to estimate the effect parameter(s) without any consideration of the hazard function. This allows one to analyse the data to ascertain the survival probability. The subsequent sub-section will now explain how the above model will be applied in this research.

4.6 Model Application

This research utilises the methodology described above to examine the impact of a number of explanatory variables on customer retention behaviour. Prior research using the Cox PH model to understand customer retention behaviour, has been undertaken by Van den Poel and Larivière (2004). However, the application of the Cox PH model in this thesis departs from this previous research in several important ways.

First, the research consolidates three sources of data: customer complaints, customer transactions and customer characteristics and to obtain a more complete view of behaviour. Also, in contrast to prior work it adds *complaint variables* in the examination of customer retention behaviour. Second, it uses *a longitudinal design* to allow for both time varying and time invariant explanatory variables within a semi-parametric model. This permits the inclusion of unobserved heterogeneity involving multiple customer records. Third, the research *compares customer retention behaviour of two products experiencing different product life cycle* - growing and declining stages. Specifically, this analysis serves to provide an understanding of those factors that are common across two products and those factors that vary between them.

Following the identification of the explanatory variables discussed in Chapter 2 and 3, a hierarchical modelling approach will be applied. This approach is appropriate because of the longitudinal analysis proposed in this study, of which there are repeated measurements from many customers over time (Singer & Willett, 2003; Snijders, 2011). Furthermore, prior research such as Reinartz and Kumar (2003), Nitzan and Libai (2011) utilised this approach in examining effects of several explanatory variables.

In this study, three models for each product category are proposed to estimate customer retention behaviour. First, the baseline model estimates customer retention behaviour as a function of the four identified customer characteristic variables and previously included variable, usage. A baseline model is often used as an entry criterion especially for studies dealing with large number of explanatory variables (Nitzan & Libai, 2011). One such benefit of a base model is to gain structural form of the parameter estimates for intuitive explanation. Furthermore, including selected variables at the beginning of the modelling process ensures that the parameter estimates are not biased especially when variable is regressed on another (Cox & Wermuth, 1996).

Second, a comprehensive model is further proposed to include explanatory variables that were derived from the exhaustive search of literatures covered in number of studies. Although a baseline form is good for parameter estimation, but it is limited in terms of interpreting customer retention behaviour. Since the aim of this research is to examine customer retention behaviour, the comprehensive model helps to explore additive effect of focal explanatory variables. In addition, this form of model helps to make inferences about each of the effect that leads to a more complete understanding of customer retention behaviour.

Finally, the third model proposed in this study examines all possible interaction based on the information from the extant literature. According to Brambor, Clark and Golder (2006) examining interaction is part of the modelling process to understand how much the interactions could explain above and over the main effects. This research aims to report the effect of interaction of possible explanatory variables actually have on the customer retention behaviour; in a way that is indifferent to whether the interaction is significant, or even present in the model. This would help to produce clear and interpretable results that further refine understanding on customer retention behaviour.

Each model described above provides differing level of insights but all three proceeds into comprehensive understanding of customer retention behaviour. The following Sub-section specifies the list of variables included in the three models using the model equation proposed in Section 4.3 earlier.

4.6.1 *Baseline Model*

A baseline model is estimated using customer characteristic and previously included variable. This consists of age, gender, household income, household size and usage and separate models are proposed for customers subscribing to voice and Internet services. The aim of the baseline model is to examine some of the previously included variables in the literature. From this model, we will confirm the relationship between the variables and customer defection behaviour from the previous modelling exercise. Using equation number (4) in Section 4.3, the baseline model is expressed as follows:

$$
\begin{gathered}
h_i(t) = h_0(t)\exp[\beta_1 * \mathrm{Gender(Female)_i} + \\
\beta_2 * \mathrm{IncomeGroup(Medium)_i} + \beta_3 * \mathrm{IncomeGroup(High)_i} + \\
\beta_4 * \mathrm{HouseholdSize(3-4)_i} + \beta_5 * \mathrm{HouseholdSize(5-6)_i} + \\
\beta_6 * \mathrm{HouseholdSize(>6)_i} + \beta_7 * \mathrm{AgeGroup(26-40)_i} + \\
\beta_8 * \mathrm{AgeGroup(41-55)_i} + \beta_9 * \mathrm{AgeGroup(>55)_i} + \\
\beta_{10} * \mathrm{Usage_i} + \gamma_i]
\end{gathered}
$$

Where:

$h_i(\mathrm{t})$ is the hazard of churning for customer i at time t.

$h_0(\mathrm{t})$ is the baseline hazard of churning.

γ_i is the random effect for customer i. These random components are assumed to be independently and identically distributed as normal random variables with mean 0 and unknown variance.

4.6.2 *Comprehensive Model*

However, the aim of this research is to extend beyond the current knowledge of customer retention. Thus, a comprehensive model is estimated comprising of all the explanatory variables discussed in Chapters 2 and 3. Specifically, the effect of customer complaints, customer transactions and customer characteristics variables will be examined for customers subscribing to voice and Internet services. Each variable included in the model is assumed to effect customer defection behaviour directly rather than the intervening of other variables in the model. By comparing the baseline to a comprehensive model, we will gain insights from a richer form of model. It will reveal factors that increases or decreases customer defection behaviour. In addition, this model will highlight importance of adding more explanatory variables to capture the holistic effect on customer defection behaviour overtime. Using equation number (4) in Section 4.3, the comprehensive model is expressed as follows:

$$
\begin{gathered}
h_i(t) = h_0(t)\exp[\beta_1 * \mathrm{Gender(Female)_i} + \\
\beta_2 * \mathrm{IncomeGroup(Medium)_i} + \beta_3 * \mathrm{IncomeGroup(High)_i} + \\
\beta_4 * \mathrm{HouseholdSize}(3-4)_\mathrm{i} + \beta_5 * \mathrm{HouseholdSize}(5-6)_\mathrm{i} + \\
\beta_6 * \mathrm{HouseholdSize}(>6)_\mathrm{i} + \beta_7 * \mathrm{AgeGroup}(26-40)_\mathrm{i} + \\
\beta_8 * \mathrm{AgeGroup}(41-55)_\mathrm{i} + \beta_9 * \mathrm{AgeGroup}(>55)_\mathrm{i} + \\
\beta_{10} * \mathrm{Usage_i} + \\
\beta_{11} * \mathrm{LengthofRelationship_i(t)} + \\
\beta_{12} * \mathrm{NumberofCommunication_i(t)} + \beta_{13} * \mathrm{ComplaintIdentifier(complainants)_i(t)} \\
\beta_{14} * \mathrm{NumberofComplaints_i(t)} + \\
\beta_{15} * \mathrm{TimeSinceLastComplaints_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{16} * \mathrm{Severity(Low)_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{17} * \mathrm{Severity(High)_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{18} * \mathrm{Lengthofrecovery_i(t)} + \gamma_\mathrm{i}]
\end{gathered}
$$

Where:

h_i(t) is the hazard of churning for customer i at time t.

h_0(t) is the baseline hazard of churning.

all variables with (t) are time varying, i.e. their values change over time. In particular the values of these variables change over 36 months (or 36 rows of records for one customer).

γ_i is the random effect for customer i. These random components are assumed to be independently and identically distributed as normal random variables with mean 0 and unknown variance.

4.6.3 *Comprehensive Model with Interaction*[23]

Often in customer retention studies, the effect of customer defection behaviour with respect to an explanatory variable to depend on the magnitude of yet another explanatory

[23] Following the approach suggested by Cox and Wermuth (1996), this research examined possible interactions among all the explanatory variables for both the voice and Internet products. This includes (a) fitting and testing all interaction terms, one at a time, and (b) plot their corresponding p-values. The results were found statistically significant for the combination: usage and age for Internet; length of recovery and

variable (Keaveney, 1995). Thus, building upon the comprehensive model, interaction amongst all the variables on customer defection behaviour is further evaluated for customers subscribing to voice and Internet services. One benefit of including interaction effects in this study is to capture the casual effects of interacting explanatory variables on customer defection behaviour. This could suggest that customer make an overall global assessment rather than weighing on each of the factors individually. Using equation number (4) in Section 4.3, the interaction model is expressed as follows:

$$
\begin{gathered}
h_i(t) = h_0(t)\beta_1 * \mathrm{Gender(Female)_i} + \\
\beta_2 * \mathrm{IncomeGroup(Medium)_i} + \beta_3 * \mathrm{IncomeGroup(High)_i} + \\
\beta_4 * \mathrm{HouseholdSize}(3-4)_\mathrm{i} + \beta_5 * \mathrm{HouseholdSize}(5-6)_\mathrm{i} + \\
\beta_6 * \mathrm{HouseholdSize}(>6)_\mathrm{i} + \beta_7 * \mathrm{AgeGroup}(26-40)_\mathrm{i} + \\
\beta_8 * \mathrm{AgeGroup}(41-55)_\mathrm{i} + \beta_9 * \mathrm{AgeGroup}(>55)_\mathrm{i} + \\
\beta_{10} * \mathrm{Usage_i} + \\
\beta_{11} * \mathrm{LengthofRelationship_i(t)} + \\
\beta_{12} * \mathrm{NumberofCommunication_i(t)} + \beta_{13} * \mathrm{ComplaintIdentifier(complainants)_i(t)} \\
\beta_{14} * \mathrm{NumberofComplaints_i(t)} + \\
\beta_{15} * \mathrm{TimeSinceLastComplaints_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{16} * \mathrm{Severity(Low)_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{17} * \mathrm{Severity(High)_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{18} * \mathrm{Lengthofrecovery_i(t)} + \\
\beta_{19} * \mathrm{LengthofRelationship_i(t)} * \mathrm{Severity(High)_i(t)} * \mathrm{ComplaintIdentifier(complaints)_i(t)} + \\
\beta_{20} * \mathrm{IncomeGroup(High)_i} * \mathrm{AgeGroup}(26-40)_\mathrm{i} + \\
\beta_{21} * \mathrm{IncomeGroup(High)_i} * \mathrm{AgeGroup}(41-55)_\mathrm{i} + \\
\beta_{22} * \mathrm{IncomeGroup(High)_i} * \mathrm{AgeGroup}(>55)_\mathrm{i} + \\
\beta_{23} * \mathrm{IncomeGroup(Medium)_i} * \mathrm{AgeGroup}(26-40)_\mathrm{i} + \\
\beta_{24} * \mathrm{IncomeGroup(Medium)_i} * \mathrm{AgeGroup}(41-55)_\mathrm{i} + \\
\beta_{25} * \mathrm{IncomeGroup(Medium)_i} * \mathrm{AgeGroup}(>55)_\mathrm{i} + \\
\beta_{26} * \mathrm{Usage_i} * \mathrm{AgeGroup}(26-40)_\mathrm{i} + \\
\beta_{27} * \mathrm{Usage_i} * \mathrm{AgeGroup}(41-55)_\mathrm{i} + \\
\beta_{28} * \mathrm{Usage_i} * \mathrm{AgeGroup}(>55)_\mathrm{i} + \gamma_\mathrm{i}]
\end{gathered}
$$

---

severity (high) for both Internet and voice; age and income for both voice and Internet. Therefore, only these significant variables were included in the interaction models and are reflected in the equation term.

Where:

$h_i(\mathrm{t})$ is the hazard of churning for customer i at time t.

$h_0(\mathrm{t})$ is the baseline hazard of churning.

γ_{i} is the random effect for customer i. These random compoments are assumed to be independently and identically distributed as normal random variables with mean 0 and unknown variance.

4.7 Summary

This Chapter discussed available methods to model customer retention behaviour. The Cox Proportional Hazard (PH) model is found to be the model most suitable for managing the occurrence of censoring and the time varying explanatory variables prevalent in this study. The analytic framework outlined in this chapter provides two primary advantages over other methods: (i) it offers richer insights regarding the effects of explanatory variables on customer retention behaviour; and, (ii) it correctly estimates customer survival probability.

Three models proposed in this research includes (i) baseline model- a simpler form of model to capture rudimentary effects of customer characteristic and usage on customer retention behaviour; (ii) comprehensive models- including all focal explanatory variables identified in the literature review with aim to obtain incremental effect on customer retention behaviour; (iii) interaction among all the focal variables to further assess the confounding effects and to report how much the interaction explains over the results found in the comprehensive model.

Next Chapter which follows will present the evaluations among the three proposed models in line with the analysis procedure covered in this Chapter. These include comparing survival probability for customers subscribing to two products- voice and Internet; evaluation on the proportional assumption; estimation of three models using appropriate statistical package and evaluation on the plausible model given the research aim. Finally the results of the analyses are presented, evaluated and discussed.

CHAPTER 5: ANALYSES AND RESULTS

5.0 Introduction

Three forms of customer retention model were estimated for two products, voice and Internet. The three models were developed in the study were: a baseline model; a comprehensive model, and an interaction model.

The *baseline model*, where a reduced number of model factors are incorporated, only contains customer characteristic and usage variables. The *comprehensive model* provides a more complete view of customer retention behaviour; it includes all complaints, transaction and customer attributes variables. In the *interactions model*, interactions among all the explanatory variables were also estimated. Three models were estimated in order to capture the incremental effect of the explanatory variables on customer retention behaviour. While the baseline model is useful to gain structural form for intuitive explanation, the comprehensive model helps to provide additive view on the effect of all the focal variables independently. Interactions model shows both the additive and interactive effects on customer retention behaviour. The final model proves to provide the best fit model for this study is the comprehensive with interaction effects. The evaluation of the model is discussed in later part of the Section.

This chapter will first briefly discuss the analysis procedure undertaken to support the model estimation. In particular, the statistical package to carry out the model estimations is discussed. Second, the survival curves for customers subscribing to voice and Internet is assessed and compared. Third, implicit in using the Cox Proportional Hazard model, the assessment of the proportionality assumption is discussed. Fourth, the selection of the final model based on the well-established evaluation criteria is delineated. Finally, the estimation results for customers subscribing to voice and Internet are presented before summarising them against each research question.

5.1 Analysis Procedure

As noted earlier, the research uses the Cox Proportional Hazard model to understand customer retention behaviour[24]. To support this aim, the data was re-structured into a customer-period with each customer having multiple records (lines of data), one per time period of observation[25]. The analysis was then carried out using SAS PHREG[26] procedure because the data for this research contains explanatory variables that are time varying and time invariant. Several studies that have used SAS to perform survival analysis are: Nitzan and Libai (2011); Reinartz and Kumar (2003) and Van den Poel and Larivière (2004). The ties[27] were handled using partial likelihood Efron approximation as discussed in the earlier Chapter 4, Section 4.4. This results in better-quality estimates in dealing with duration type data than the Berslow approximation (Bolton, 1998; Nitzan & Libai, 2011).

---

[24] An empirical check for the appropriateness of the parametric models was carried out. The Cox model was chosen as it is a robust model for fitting a wide variety of data distributions (Singer & Willett, 2003).

[25] The information is based on Singer and Willett (2003).

[26] The analysis could be carried out using different statistical packages such as SPSS, SAS, Stata, R but SAS currently has the most comprehensive set of full-featured procedures for performing survival (Allison, 2010).

[27] Ties are cases in which two or more customers have an exact defection time. It is possible to correct this type of bias with an appropriate method which was discussed in Chapter 4, Section 4.4.

Following this analysis procedure, the study now considers the survival distribution for customers subscribing to voice and Internet services. The section below provides an overview of customer's survival probability as a function of time.

5.2 Survival Charts

As highlighted in Chapter 4, Section 4.2.5 understanding the survival distribution is one of the building blocks in performing survival analysis. It is a first step of univariate analysis of customer's survival probability as a function of time. One of the widely used approaches is the Kaplan-Meier method[28], a non-parametric based descriptive statistics for survival data (Allison, 2010; Mills, 2011). It is important to note that although mean is generally used to describe the central tendency of a distribution, this is not the case in understanding the survival distribution. In such cases the median is often preferred because a small number of short or long lifetime individuals would cause the mean survival time to be disproportionately large or small (Reinartz & Kumar, 2003). The median survival time is defined to be the time at which the survival curve crosses 50 per cent survival probability (Allison, 2010; Mills, 2011). In dealing with actual customer behavioural data as in this study, it frequently happens that more than 50 per cent of observations in the data are censored[29]; therefore the median is unspecified. In this type

[28] The survival curves are obtained from the product limit survival table. Due to the size of the observations, it is not feasible to append them.

[29] Censored is a term used to indicate that the customer has not experienced the event at the end of the observation period, so its true duration of their survival is unknown. The concept has been covered sufficiently in Chapter 4, Section 4.1.3.

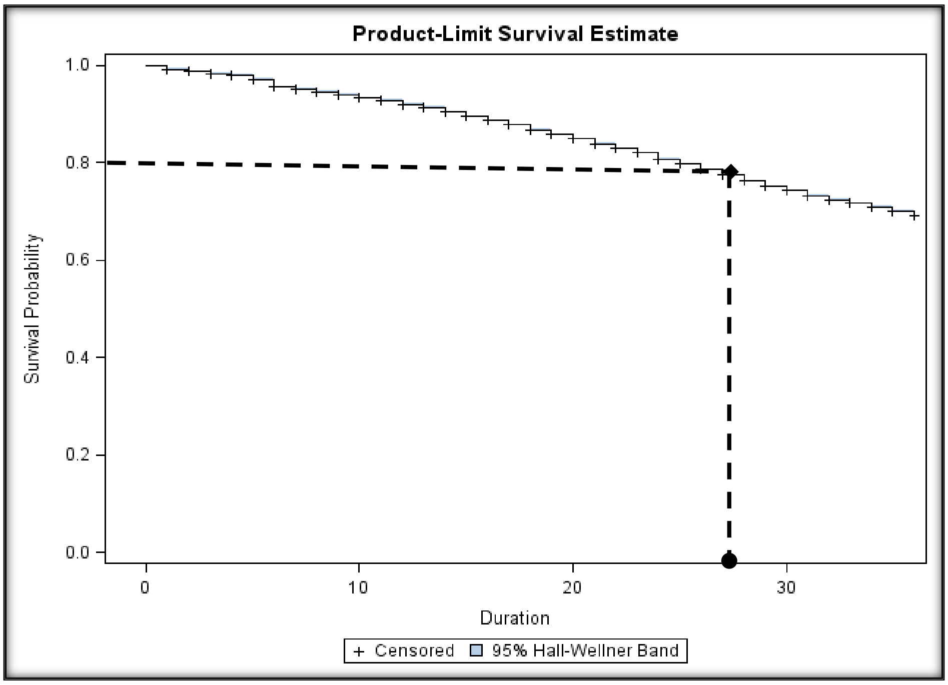
of situation another quantile such as 80 per cent is recommended (Allison, 2010; Mills, 2011).

Application of this method to estimate survival probability for voice subscription customers is represented in Figure 5.1 below. Similarly, Figure 5.2 plots Internet subscription customers' survival probability. For both Figures 5.1 and 5.2 below, the "*X-axis*" represents *duration* of the observation period measured in months. As highlighted in Chapter 3, Section 3.4, the study uses the monthly data from 1st January 2010 to 31st December 2012 and therefore the window of observation is for 36 months. While, "*Y-axis*" is represented by *survival probability* with boundaries at 0.2, 0.4, 0.6 up till 1.0. As noted, censored observation represents information where the customer's survival time is incomplete at the end of the study duration. This represented 77 per cent for voice subscribers and 75 per cent for Internet subscribers[30]. Considering that more than 50 per cent of the observations in this data are censored, a quantile of 80 per cent is used to extrapolate the survival probability for voice and Internet subscribers.

Figure 5.1 shows that for customers subscribing to voice services the survival probability at 80 per cent quantile (black dotted line) is approximately 27 months. The likelihood of customer defection is lower at the beginning of the observation period but as the time passes the slope becomes much steeper and the plots eventually reaches a plateau towards the end of the observation period.

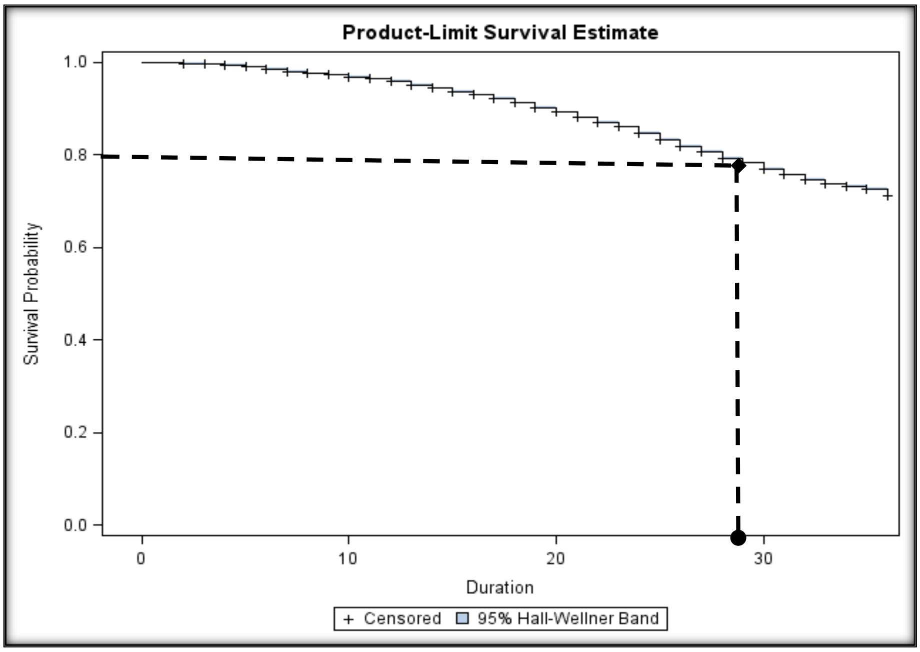
[30] The censoring information is gathered from the final table of output from running the PHREG operation.

Figure 5.1: Survival Estimates for Voice Services



Likewise, for customers subscribing the Internet service, depicted in Figure 5.2, the survival probability at 80 per cent quantile (black dotted line) is approximately 29 months. The likelihood of customer defection is lower at the beginning of the observation period but the decreasing shape suggests that their survival probability becomes lower with longer duration of time.

Figure 5.2: **Survival Estimates for Internet Services**



From this perspective, the results contradict previous findings that customers with longer tenure are less susceptible to the risk of defection (Reinartz & Kumar, 2003; Schweidel et al., 2008). These observations could be explained by the fact that customers in this research have varied lengths of relationship with the firm, thus the defection percentage is lower at the beginning and then increases through the later part of the duration of the relationship.

The decision to include the varied tenure in the model is to capture the effect of length of relationship on customer retention behaviour. Further, including this type of information could provide significant managerial implications for the firm to consider with the risk of defection varying based on the length of time a customer has been with the firm.

Between the two products, voice and Internet, the length of duration for customers subscribing to Internet services (approximately 29 months) appears slightly longer than those of voice services (approximately 27 months). It is obvious from the plots that the steepness of the curve is more prominent in voice than in Internet service, indicating the risk of defection is likely to be higher for voice. Although the survival in the two product groups varies, it is impossible to make a conclusive finding without additional information. Thereby, to further understand customer retention behaviour, the study included explanatory variables, as discussed in Chapters 2 and 3 in the model estimation. However, before outlining the model estimation, the testing of the proportional assumption is considered next.

5.3 Testing the proportionality assumptions

As discussed in Chapter 5 Sub-section 5.4, implicit in using the Cox PH model, is the proportionality assessment. The residuals based Schoenfeld eliminations are meant to test if there is a correlation between time and the amount of error that is displayed across it. If there is any temporal correlation, it implies the presence of temporal autocorrelation. In that case, the PH assumption is violated and the results of the analysis cannot be interpreted because the effect of the independent variable on the dependent variable is seen to vary with time[31]. If interactions significantly exist, the proportional assumption is violated (Cox, 1972).

To support this assessment, a graphical examination of the Schoenfeld residuals was conducted using SAS functions. This method takes less time and is robust considering

[31] Evaluation of proportional hazards is based on Schoenfeld residuals (SR) plot. The SR plot for a given continuous variable should not show deviation from a straight line but remains flat. Brief discussion is provided in Appendix C.

the large sample size deployed in this research (Grambsch & Therneau, 1994). The assessment is usually carried out for time varying explanatory variables: communication encounter, number of complaints, length of recovery, length of relationship, usage and time since last complaint.

The graphical method consists in creating a scatter plot of partial residuals of the explanatory variable as y axis and time-to event as the x variable. The results of this check indicate that there is no relation between the residuals and survival time and therefore the time varying explanatory variables in the model meet the proportionality assumptions. These figures are presented in Appendix C.

Having completed an assessment of assumptions that is considered reasonable, the following sub-section presents the findings and interpretation from the models which are used to investigate the retention behaviour for customers subscribing to voice and Internet services.

5.4 Findings

Three forms of model were estimated for each product, voice and Internet services. Complete results for all the three models estimation are presented in Appendix D. Of the three models estimated, comprehensive model with interaction terms was the final model for both voice and Internet. This result is presented in Table 5.1 for voice subscribers and Table 5.2 for Internet subscribers. The information from these tables is now summarised.

Table 5.1: Summary Result for Voice Subscribers

Row Number	Explanatory Variables	Hazard Ratio	Risk of Defection	Effect of Retention
	Age			
	Group 1 (19-25) (reference group)			
1	Group 2 (26-40)	1.027	3%	Negative
2	Group 3 (41-55)	1.330	33%	Negative
3	Group 4 (over 55)	1.405	41%	Negative
	Gender male (reference group)			
4	Female	0.994	-0.6%	Positive
	Household income			
	low (reference group)			
5	medium	1.273	27%	Negative
6	high	1.191	19%	Negative
	Household size			
	1-2 (reference group)			
7	3-4	0.927	-7%	Positive
8	5-6	0.908	-9%	Positive
9	over 6	0.919	-8%	Positive
10	Usage	1.002	2%	Negative
11	Length of relationship	0.996	-0.4%	Positive
12	Communication encounter	1.043	4%	Negative
	Camplainants (reference group)			
13	Non-complainants	2.735	175%	Negative
14	Number of complaints	0.951	-4.90%	Positive
15	Time since last complaints	0.961	-3.9%	Positive
	Severity			
	none (reference group)			
16	High	0.424	-58%	Positive
17	Low	2.989	189%	Negative
18	Length of recovery	0.994	-6%	Negative
	Income group (medium)*			
19	age (26-40)	1.196	20%	Negative
20	age (41-55)	1.062	6%	Negative
21	age (>55)	0.937	-6%	Positive
	Income group (high)*			
22	age (26-40)	2.038	104%	Negative
23	age (41-55)	1.802	80%	Negative
24	age (>55)	1.637	64%	Negative
	Length of recovery *			
25	severity (high)	1.014	1.4%	Negative
	Usage *			
26	age (26-40)	1.000	No effect	No effect
27	age (41-55)	1.000	No effect	No effect
28	age (>55)	1.000	No effect	No effect

Note:

1. All variables are statistically at $p<0.0001$ except for gender and age group 2
2. The standard error in this model is reported in Appendix E1. All the values are below than 2.0 that points towards a good fit model.

Table 5.2: Summary Result for Internet Subscribers

Row Number	Explanatory Variables	Hazard Ratio	Risk of Defection	Effect of Retention
	Age			
	Group 1 (19-25) (reference group)			
1	Group 2 (26-40)	1.468	47%	Negative
2	Group 3 (41-55)	1.725	73%	Negative
3	Group 4 (over 55)	1.697	70%	Negative
	Gender male (reference group)			
4	Female	0.886	-11%	Positive
	Household income			
	low (reference group)			
5	medium	1.462	46%	Negative
6	high	1.623	63%	Negative
	Household size			
	1-2 (reference group)			
7	3-4	0.940	-6%	Positive
8	5-6	0.876	-12.40%	Positive
9	over 6	0.942	-5.8%	Positive
10	Usage	0.996	-0.4%	Positive
11	Length of relationship	0.998	-0.2%	Positive
12	Communication encounter	1.022	2.2%	Negative
	Camplainants (reference group)			
13	Non Complainants	3.862	287%	Negative
14	Number of complaints	0.843	-15.70%	Positive
15	Time since last complaints	0.976	-2.4%	Positive
	Severity			
	none (reference group)			
16	High	0.351	-70%	Positive
17	Low	4.745	375%	Negative
18	Length of recovery	1.004	0.40%	Negative
	Income group (medium)*			
19	age (26-40)	0.773	-23%	Positive
20	age (41-55)	0.797	-20%	Positive
21	age (>55)	0.896	-10%	Positive
	Income group (high)*			
22	age (26-40)	1.278	28%	Negative
23	age (41-55)	1.408	21%	Negative
24	age (>55)	1.516	52%	Negative
	Length of recovery *			
25	severity (high)	1.011	1.1%	Negative
	Usage *			
26	age (26-40)	0.999	-0.10%	Positive
27	age (41-55)	0.999	-0.10%	Positive
28	age (>55)	0.999	-0.01%	Positive

Note:

1. All variables are statistically significant at $p<0.0001$.
2. The standard error in this model is reported in Appendix E2. All the values below than 2.0 points towards a good fit model.

On the left-hand side of Tables 5.1 and 5.2, row numbers linking thirteen explanatory variables together with their categories are listed. The results for categorical variables will be compared to the reference group. The following three columns to the right-hand side of these tables provide details the hazard ratio[32], the risk of defection[33] in per cent and the effect of retention[34].

5.4.1 *Best Model Fit*

In this study, based on the three models presented, the final model was assessed based on a well-established evaluation criteria discussed in Appendix F. This information is used to choose the best fit between the models because of the limitations in the ordinary statistical tests. The results of the assessments are provided in Appendix F. From this assessment, it is found that the best fitting model for both voice and Internet services is comprehensive with interaction effects. Models that produce smaller values[35] indicates strong evidence in terms of information loss and therefore will be selected as a best fit model (Akaike, 1981; Schwarz, 1978). The results indicate that there is marginal difference in terms of information loss between the comprehensive and the interaction models for both voice and Internet services. However, even small increments between

---

[32] Hazard ratio describes *increase* or *decrease* of risk associated with a unit of increase in the variable.

[33] Risk of defection describes the rate of risk in per cent. As an example, for variable usage (row number 10), a hazard ratio of 1.002 means that the likelihood of defection is 0.2 per cent higher at any given time point studied.

[34] The effect of retention is identified by making inferences from the hazard ratio. A value of more than 1 implies that the risk of defection will increase thus a negative effect on retention. While, a hazard ratio of less than 1 indicates risk of defection is lower and therefore have positive effect on retention.

[35] Value here refers to the Akaike's Information Criterion (AIC) or Schwarz-Bayes Criterion (SBC) (Akaike, 1981; Schwarz, 1978) and these information is obtained from running the PHREG operation in SAS.

these models imply a better fit given the penalized-likelihood criteria during the model evaluation (Singer & Willet, 2003).

5.4.2 *Significance of Variables*

Finally, the notes at the bottom of each table provide details of the significance of explanatory variables (the p-value). Central to many different analysis techniques, a small p-value indicates that the difference is larger than can reasonably be explained as a chance occurrence (Gelman & Stern, 2006). Thereby, it provides sufficient evidence to conclude that the relationships observed between variables can be interpreted as being true. In this study for customers subscribing to voice services all variables, except for gender and age group within the category of 26 to 40 years old are found to be statistically significance. For customer subscribing to Internet services, all the variables are found to be statistically significant. Furthermore, in examining all possible interaction between the explanatory variables36, several were found to be statistically significant. These are reported to Tables 5.1 and 5.2 and will be discussed later in the Section 5.5 in the thesis.

5.4.3 *Research Questions and Results*

This section is divided into eight sub-sections, each of which presents the results relating to the research questions. In understanding the retention behaviour, comparison is made between customer subscribing to voice and Internet services. As highlighted

[36] This research examined all possible interactions among all the explanatory variables for both the voice and Internet services. The results found statistical significant for the combination: usage and age for Internet; length of recovery and severity (high) for both Internet and voice; age and income for both voice and Internet. Therefore, only these significant variables were included in the interaction models.

earlier, Table 5.1 shows the estimated result for voice subscription customers and Table 5.2 shows the estimated result for Internet subscription customers.

5.4.3.1 Research Question 1: Can customer complaining behaviour explain customer retention behaviour?

In Section 2.4.1, the parameter estimate is expected to be positively related to customer retention behaviour for complainants, as the firm would be expected to have greater opportunity to resolve the issues.

Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for non-complaints as compared to the complainants (Row 13) is 2.735. Table 5.2 shows the estimated result for customers subscribing to Internet services. The hazard ratio for non-complainants as compared to the complainants in Table 5.2 is 3.862. These results for voice and Internet services shows positive parameter estimates with the results significant at $p<.0001$.

From the results, it can be deduced that a customer who never complains is at a greater risk to defection than one who has ever submitted a complaint. For voice services, the risk of defection is 175.0 per cent while customers subscribing to Internet services the risk of defection are 287.0 per cent. Although the hazard ratio is much smaller for voice service, it is observed that there is positive effect on the hazard attributable to being a complainant. The effect is more pronounced for Internet customers than for voice customers.

The results parallels the findings of Lariviere and Van den Poel (2005) and Ahn et al. (2006) that a customer who lodges a complaint is overall "more involved" or "more active" with the firm, and therefore less likely overall to defect. Therefore, customers who complain, have hopes that the firm will make satisfactory improvement or enhancement to the quality of service. Hence, one interpretation would be that one way to please the customer, and minimise defections, would be to manage their complaints diligently.

For customers within the non-complainants group, there appears to be two latent cohorts of customers. Specifically, those customers satisfied with the quality of service; and then those customers who may be dissatisfied with the service. This particular dissatisfied non-complainant may simply defect having never lodged a complaint as a result of their unaddressed dissatisfaction. However, unavailability of suitable firm data prevented this research deriving conclusive findings regarding this dissatisfied non complainant cohort. At the same time, collecting this type of would likely help the firm to isolate and address the defection threat of these customers who may not complaint regardless of their dissatisfaction.

5.4.3.2 Research Question 2: Does the number of complaints explain customer retention behaviour?

As discussed in Section 2.4.1 it is likely that number of complaints is associated with customer retention but the direction of the response parameter for number of complaints is uncertain. That is, the response parameter could be negatively or positively related to customer retention.

Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for number of complaints (Row 14) is 0.951. Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for number of complaints (Row 14) is 0.843. From the result, customer subscribing voice and Internet services shows a hazard ratio smaller than 1 and the results are significant at $p<.0001$.

From the results, it is deduced that the risk of defection is 5.0 per cent lower for voice services when there is at least one complaint compared with no complaints at all. While, the risk of defection for the Internet customers is 15.7 per cent lower when there is one or more complaints than when there's no complaints.

Between the two products, risk of defection is more pronounced for voice rather than Internet services. One possible explanation is that the importance of Internet as opposed to the voice services has become an undisputable fact. Hence, reliability and reliance on Internet as an essential service could surpass the voice service. As noted by Rappa (2004) growth of information and communication has made Internet service a mainstream activity. Customer's dependency on the Internet services could be the impetus for more calls to the customer care centre. That is, the more complaints a customer submits, the more engaged he or she is with the product, and less likely to defect, if their complaints are addressed by the firm. Indeed, if a customer goes through the trouble of contacting the firm and the firm promptly addresses the customer's complaint, the customer is more likely to stay with the firm. Therefore, the results of the models support the proposition that number of complaints is positively associated with retention.

5.4.3.3 Research Question 3: How does "recency" of complaints impact customer retention behaviour?

In Section 2.4.1, the response parameter for time since the last complaint is expected to positively relate to customer retention. That is, a customer who has recently made a complaint is more likely to defect than one who complained earlier.

Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for time since last complaints (Row 15) is 0.961. Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for time since last complaints (15) is 0.976. From the result, customer subscribing voice and Internet services shows a hazard ratio smaller than 1 and the results are significant at $p<.0001$.

Overall, after controlling for all of the other effects related to complaints, it can be deduced that a customer's risk of defection decreases as time passes since they last complained. The risk to defection decreases by 4 per cent for voice service customers and 2.4 per cent for Internet service customers for each month that passes since their last complaint.

When considered in combination with the last set of provisional conclusions, this suggests that a customer who complains is overall less likely to defect whereas, for a customer who lodges multiple complaints within a short period of time, the positive effect of "recency" of complaint is reduced. This makes sense in so far as having no complaints for a long time may be an indicator of improved quality of service and hence, customer satisfaction. One aspect requiring consideration in this context is survival bias. For example, customers wishing to change provider may have already defected after their initial complaint. And customers with limited options may remain for a longer time despite their dissatisfaction and complaints.

The results also show the hazard of defection being more pronounced for Internet (4.0 per cent) than voice (2.4 per cent). This could be due to Internet industry consolidation and therefore availability of improved services. Voice services have been in operation much longer than Internet services and voice service is in its declining stage. At the same time, there is support for the view that Internet services, in general, have higher defection rates (Blattberg et al., 2008).

5.4.3.4 Research Question 4: Does the severity of complaints have greater effect on customer retention behaviour?

As discussed in Section 2.4.1, the explanatory variable severity is associated with customer retention but the direction of this explanatory variable is uncertain as previous findings are not conclusive.

Table 5.1 shows the estimated results for customer subscribing voice services. The hazard ratio comparing high severity to none (Row 16) is 0.424 while the hazard ratio comparing low severity to none (Row 17) is 2.989. Table 5.2 shows the estimated result for customer subscribing to Internet services. The hazard ratio comparing high severity to none (Row 17) is 0.351; while the hazard ratio comparing low severity to none (Row 17) is 4.745.

The results indicate the difference in outcomes of customer complaints categorized as high severity versus low severity. Low-severity complainants are a greater risk to defection than non-complainants, with defection rates of 189.0 per cent and 375.0 per cent for voice and Internet products respectively. However, high-severity complainants are less likely to defect compared with non-complainants, with lower likelihood to defection of 58.0 per cent for voice and 70.0 per cent for Internet products.

The risk of a customer defecting in the same month as making a complaint was greater when the complaint was of low severity. This equates a higher severity complaint with a greater level of justification for the complaint explains this somewhat counter-intuitive result as a consequence of recovery success. The findings can be best explained by the service recovery paradox which posits that it is likely that customers receiving good service recovery may be more satisfied than prior to the service failure itself (De Matos, Henrique, & Rossi, 2007). Because of much higher priority being given to more severe service failures by Firm A, the service recovery paradox provides an explanation as to why more severe failure results in less likelihood of defection. Lariviere and Van den Poel (2005) made similar inferences that due to greater level of justification for high severity complaints explain the counter intuitive results.

Further, a simple cross tabulation analysis for voice services customers experiencing incidents categorized as high severity found that 84 per cent remained with the firm while 16 per cent defected. Of incidents categorized as low severity, 72.0 per cent remained with the firm while 28 per cent defected. Internet services analysis resulted in similar high severity findings (86 per cent remained and 14.0 per cent defected) and low severity findings (75.0 per cent remained and 25.0 per cent defected). The high percentage of retention across the two product categories explains this somewhat counter-intuitive result as a consequence of recovery success.

Some complaints are complex and take time to resolve. Discussion with executives at Firm A revealed that complaints categorised as high severity includes constant technical problems relating to fault repairs, network coverage, slow download or upload, reliability of the connection, billing and payment which cannot be resolved promptly and requires longer resolution time. In such cases, the firm disclosed that more resources are committed and the complaints were systematically recorded and followed up to a resolution. Related to this, unresolved severe cases are turned over to the employees with higher level of experience and expertise. The higher priority vis á vis escalation to experienced employee and action to resolve the unceasing problem, helps redress customer complaints; and therefore customer perceptions are dominated on how the complaints were treated more than the problems itself (Homburg & Fürst, 2005). This helps to explain the link between severe cases and customer retention behaviour.

5.4.3.5 Research Question 5: How does the length of recovery influence customer retention behaviour?

As discussed in Section 2.4.1, it is expected that the response parameter for 'length of recovery' is expected to be positively related to customer retention if the firm is successful in its recovery efforts.

Table 5.1 shows the estimated result for customers subscribing to voice services. The hazard ratio for length of recovery (Row 18) is 0.994. Table 5.2 shows the results for customer subscribing Internet services. The hazard ratio for length of recovery (Row 18) is 1.004. The hazard ratio is smaller than 1for voice and significant at $p<.0001$. While the hazard ratio is greater than 1for Internet services and the results are significant at $p= 0.05$.

From the results, it can be inferred that the risk of defection decreases by 6.0 per cent and increases by 0.40 per cent, for every extra day the firm takes to recover from the failure, for voice and Internet failure respectively. The results show that a longer recovery time for a customer's complaint, increases the risk to defection in the same month. Since the recovery time from a failure with the firm reflects additional time and other cost to users, it is expected to be a factor influencing customer retention behaviour. This implies time taken to resolve a complaint is a good indicator of the firm's ability to handle complaints that effect customer retention behaviour. Further, recovery time has been identified as an important criterion in complaint satisfaction. This is because a customer's evaluation of and satisfaction with the firm's recovery effort (as opposed to the failure itself) may have significant impact on retention behaviour and defections (Blodgett et al., 1997; Clemmer & Schneider, 1996; Tax & Brown, 1998)

The risk of defection for voice is much lower than that for Internet services. There are possible explanations for this contrasting defection behaviour. This effect may be attributable in part that Internet subscribers generally make fewer complaints (79.0 per cent) than voice subscribers (70.0 per cent)[37]. As a consequence of this, it confirms that non-complainants have greater likelihood of defection given the findings from research questions 1.

[37] The information is obtained from Section 3.7, Table 3.3: Descriptive statistics for categorical variable.

Sub-research questions relating to customer transaction explanatory variables

5.4.3.6 Research Question 6: Does communication with customer support help or damage customer retention?

In Section 2.5, parameter estimates is expected to be positively related to customer retention as a greater number of communication encounters between the customer and firm will lead to higher retention.

Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for 'communication encounter' (Row 12) is 1.043. Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for 'communication encounter' (Row 12) is 1.022. In specific terms, the hazard ratio are greater than 1 for voice and Internet services and both significant at $p<.0001$.

The results suggest that the total number of overall communication encounters increases a customer's risk of defection. Specifically, risk of defection increases by 4.0 per cent and 2.2 per cent respectively for voice and Internet services for every number of communications with the firm. The results are significant and somewhat surprising in that no support was found for the parameter estimations in the direction expected.

It may be that customer complaints (e.g., service failure, voice quality) have far greater impact than communication encountering incidents such as inquiry, feedback and customer request. The results could be explained from findings by Bitner (1990) that some key interactions between customer and firm can have relatively greater impact than other actions, which are viewed as having lesser impact. The results may reflect that the firm's communication effort and responsiveness to the enquiries play a crucial role in shaping customer-firm encounters.

5.4.3.7 Research Question 7: Does length of relationship and usage supports customer retention behaviour over time?

Both usage (Row 10) and length of relationship (Row 11) and are discussed in this section. In Section 2.5, the parameter estimate for explanatory variable 'length of relationship' is expected to be positively related to customer retention. That is, customers who have had longer relationships will remain with the firm. And, the explanatory variable 'usage' is expected to be positively related to retention.

Length of relationship:

Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for length of relationship (Row 11) is 0.996. Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for length of relationship (Row 11) is 0.998. The hazard ratio are less than 1 for voice and Internet services and both are significant at $p<.0001$.

The results suggest that for both the voice and Internet products, an additional month of past relationship corresponds to 0.2 per cent and 0.4 per cent reduction in the risk of defection. The result suggests an overall customer approval of the service. When a customer stays longer with the firm this signifies high level of underlying involvement and overall utility of the services. This finding reinforces the importance of length of relationship with the firm we find supports previous research (Bolton, 1998; Schweidel et al., 2008; Wirtz et al., 2014). Likewise, a customer unsatisfied with their service utility or quality would most likely terminate their relationship at an earlier stage. The findings are crucial in this research since customers included in the models are having non-contractual relationship and are therefore free to move to another firm without switching penalty.

The risk of defection is slightly higher for Internet than for voice. This suggests that customers subscribing to Internet services are more prone to defection than voice subscribers, perhaps indicating that Internet services benefits and utility is much higher than voice. This could be an essential precursor to enhance customer retention.

Usage:

In terms of the explanatory variable 'usage', Table 5.1 shows the estimated result for customer subscribing to voice services. The hazard ratio for usage (Row 10) is 1.002. Table 5.2 shows the estimated result for customer subscribing to Internet services. The hazard ratio for usage is 0.996. The hazard ratio is greater than 1 for voice but it is lesser than 1 for Internet services. The results are statistically significant for both models at $p<.0001$ and implies that each additional download per month corresponded to a 0.4 per cent reduction in risk of defection for Internet customers, and 2.0 per cent increase in risk of defection for voice customer.

The differences in the results for two product categories are expected. It may be that customer's behaviour could be driven by the services attributes and indication of the utility of the service to the customers. For example, the reliance of customers on Internet services coupled with the way service is provided and utilised. Unlike voice service, Internet access is offered as always available. Conversely, voice service is dependent on the access of services and customers are usually charged for local and long-distance access. This could explain customers' negative perception of the value of voice services, which, in turn, indirectly influence retention behaviour.

5.4.3.8 Research Question 8: Are retention probabilities different depending on customer's demographic characteristics?

In Section 2.6, several customer specific characteristic were examined and included in the models for both the voice and Internet services.

Age:

First, for the explanatory variable age, it is expected that the parameter estimate for older customers are positively related to customer retention. Table 5.1 shows the estimated result for customer subscribing to voice services. The reference group for voice product is Age Group 1 (19-25 years). The hazard ratio for age group 2 -Row 1 (26-40 years), comparing the reference group is 1.027. The hazard ratio comparing Group 3-Row 2 (41-55 years), to reference group is 1.33. The hazard ratio for Age Group 4-Row 3 (over 55 years), to reference group is 1.41. The hazard ratio are greater than 1 across categories and it can be deduced that the risk of defection is 2.7 per cent higher for Age Group 2, 33.0 per cent higher for Age Group 3 and 41.0 per cent higher for age group 4.

On the other hand, Table 5.2 shows the estimated result for customer subscribing to Internet services. The reference group for Internet product is age group 1 (19-25 years). The hazard ratio comparing age group 2-Row 1 (26-40 years), to the reference group is 1.47. The hazard ratio comparing age group 3-Row 2 (41-55 years), to reference group is

1.725. The hazard ratio comparing group 4-Row 3 (over 55 years), to reference group is 0.529 is 1.697. The hazard ratio is greater than 1 across categories and it can be deduced that the risk of defection is 47.0 per cent higher for age group 2, 73 per cent higher for age group 3 and 70.0 per cent higher for age group 4. All the results except for group 2 were significant at p<.0001.

Figure 5.3: Defection Rates for Age

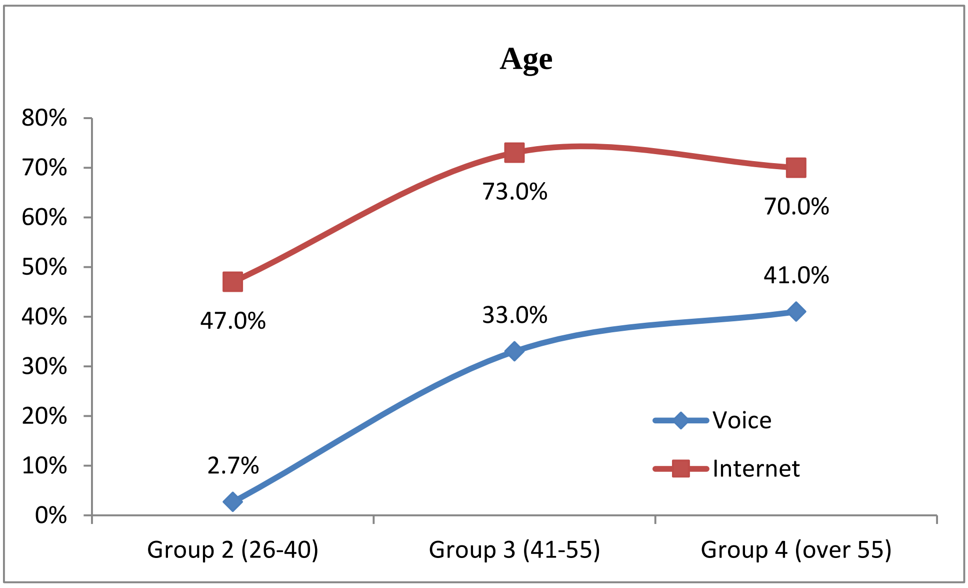


Figure 5.3 above, represents the defection rates for customers subscribing to voice and Internet services. The horizontal axis represents variable age and the categories; while vertical axis represents the defection rates.

From the graph above, risk of defection for middle-aged customers (26 - 40 years old) is 2.7 per cent higher than that of young customers (19 - 25 years old). Although this is a lot smaller than the 47.0 per cent difference observed for the Internet product, it is still significant. Older customers still have higher risk of defection compared with the younger ones: in particular, the 41 - 55-year-old is 33.0 per cent more likely to defect than the 19

- 25-year-old, and customers who are over 55 are 41.0 per cent more likely to defect. Once again, the differences are less dramatic compared with the Internet product.

For Internet services the risk of defection for middle-aged customers (26 - 40 years old) is 47.0 per cent higher than that of young customers (19 - 25 years old); older customers have even higher hazard of defecting: in particular, compared with the 19 - 25-year-old, the 41 - 55-year-old is 73.0 per cent more likely to defect, and customers who are over 55 years old are 70.0 per cent more likely to defect. All the results were significant at $p<.0001$.

Some would argue that younger people are more likely to switch telecommunication providers than their older cohort. For example, a study by the Berkeley Research Group suggests that young people are more aware of competitions amongst service providers and hence are more likely to switch. Although this research results appear counter intuitive, this finding is not an uncommon occurrence. For example, Braun and Schweidel (2011) discovered that the effect of age on customer loyalty is not monotonic. Of several possible explanations for this result, the first relates to having possibly under-sampled the younger cohort (19 - 25 years olds account for 7.0 per cent of the data). This may have inflated the higher risk of defection for older groups. This is, however, unlikely because the actual number of (2000+) observations in the younger cohort was likely large enough to prevent spurious effects of significance.

The similarity in defection rates amongst the older cohorts (especially the 41-55 and over 55 groups) suggest that older consumers are indeed more likely to defect. Apart from the younger cohort (7.0 per cent of the total sample), the other three cohorts were equally represented and each accounted for around 30 per cent of the total sample.

There are possible explanations for the likelihood of defection of older customers found in the study, including the fact that the service packages offered by the firm in this study could have been specifically designed for a younger customer cohort. This would be unsurprising as many firms expect that young people tend to churn and take careful steps to appeal to and retain their younger customers. In this scenario, older customers are not attracted to service packages geared toward younger customers and are accordingly more likely to defect. A concerted retention effort by the firm may have focused on online or email promotional offers, coupons, and discounts, and it could be easier for younger customers to access these due to their overall technological fluency.

There are a number of other potential explanations of the higher likelihood of defection by older customers. Mortality could be a potential explanation for the higher rates, especially for people over 55 years old. It is also possible that customers defected not because of switching to a different service provider, but simply due to factors outside of the control of the firm. For example, older customers may no longer make decisions for themselves, with their children or caregivers possibly in charge of decisions about voice and Internet services. Finally, a higher defection rate for middle-aged customers could be explained by lifestyle changes such as marriage, child birth and moving to a new house. These customers might have different requirements due to lifestyle changes and chose a service plan that better suited their changed needs. It is impossible to validate these possible reasons without additional data, which is beyond the scope of this research.

Gender:

Moving to the effect of the explanatory variable ‘gender’ on retention, the expectation is for this to be equivocal, and highly context dependent. Thus, the direction of the parameter estimates could be either positively or negatively related to customer retention.

For both the voice and Internet, the reference category for gender is ‘Male’. Table 5.1 shows the estimated result for customer subscribing voice services. Compared to the reference category, the hazard ratio for gender (Row 4) is 0.994. Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for gender (Row 4)’ is 0.886. The hazard ratio is lower than 1.0 for voice and Internet services, but it is significant at $p<.0001$ only for Internet service. It was not significant for voice services, thus no inferences could be drawn.

The results suggest that female customers are less likely to defect than male customers for Internet service. This difference between females and males may be caused by some aspect of firm operation such as marketing and advertising casing effects across the two gender groups. It is not possible to pinpoint the exact causes without more data and without consideration of the firm’s marketing strategy, the occupations of its customer’s base, etc. Since the effect was small to begin with, and disappears in the presence of a few additional seemingly independent variables, it may be concluded that gender is not a very good predictor of defection for these products.

Household income:

The explanatory variable ‘household income’, described in Section 2.6, is expected to be positively associated with customer retention. That is, a high income group is less sensitive to defection. The reference group for both voice and Internet services is the low income group. Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for medium income group (Row 5) compared to the reference group is 1.27. The hazard ratio for high income (Row 6) compared to reference group is 1.19. The hazard ratio is greater than 1.0 for the variable ‘household income’ and are significant at $p<.0001$ and $p=0.05$ for low and high income respectively.

Table 5.2 shows the estimated result for customer subscribing to Internet services. The hazard ratio for medium income group (Row 5) compared to the reference group is 1.46. The hazard ratio for high income (Row 6) compared to reference group is 1.62. The hazard ratio is greater than 1.0 for ‘household income’ are significant at $p<.0001$ respectively.

Figure 5.4: Defection Rates for Household Income

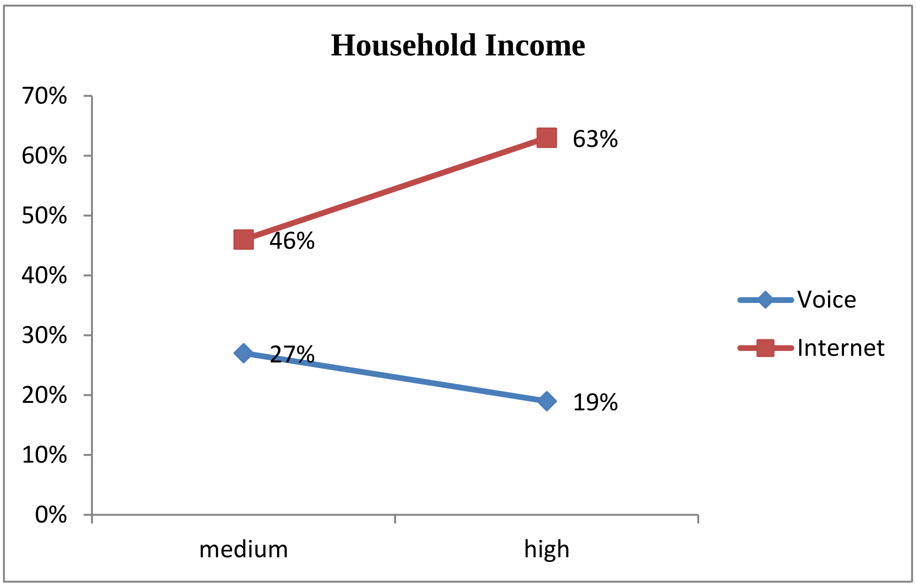


Figure 5.4 presents the defection rates for customers subscribing to voice and Internet services. The horizontal axis represents the variable 'household income' with income category and vertical axis provides information on the defection rates.

As shown in Figure 5.4, for voice subscribers the risk of defection is 27.0 per cent and 19.0 per cent higher for medium and high income group as compared to the low income group. While, for Internet subscribers the risk of defection is 42.0 per cent and 62.0 per cent higher for medium and high income group as compared to the low income group.

Across voice and Internet products, the results suggest that the risk of defection for high income earners (27.0 per cent) subscribing to the voice product compared to high income earners subscribing to Internet services (62.0 per cent). Similarly, the hazard for defecting is higher for medium income earner subscribing to Internet services (46.0 per cent) as compared to voice services (24.0 per cent). From the results it can be inferred that households with higher income tend to have a higher defection risk, for both the voice and Internet products. One possible explanation for this result is that the firm may provide more basic services at a highly competitive price; and their customers with higher incomes (and lower price sensitivity) may defect to improved service subscriptions at higher prices[38]. As a result, high-income customers are more likely to defect and low-income customers are more likely to stay with the firm. Moreover, people with higher income tend to have a higher level of education and have sufficient digital fluency to use the Internet to research competitive offers from alternative service providers.

Household size:

Finally, in Section 2.6 for the explanatory variable household size it is expected to be positively associated with customer retention as bigger number of household size would have lower likelihood to defection than smaller family size.

The reference group for both voice and Internet services is household size (1-2 people). Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for household size (3-4 people)-Row 7, compared to the reference group is 0.93. The hazard ratio for household size (4-5people)-Row 8 compared to the reference group is 0.908. The hazard ratio for household size (more than 6 people)-Row 9 compared to the reference group is 0.919. The hazard ratio is greater than 1.0 across categories and it is significant at $p < .0001$.

[38] The explanatory variable income could be acting as an indicator for certain lifestyle and the relationship between the effect of price plan and income could be examined in future research.

Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for household size (3-4 people)-Row 7 to the reference group is 0.94. The hazard ratio for household size (4-5people)-Row 8 to the reference group is 0.876. The hazard ratio for household size (more than 6 people)-Row 9 to the reference group is 0.942. The hazard ratio is greater than 1.0 across categories and it is significant at p <.0001

Figure 5.5: Defection Rates for Household Size

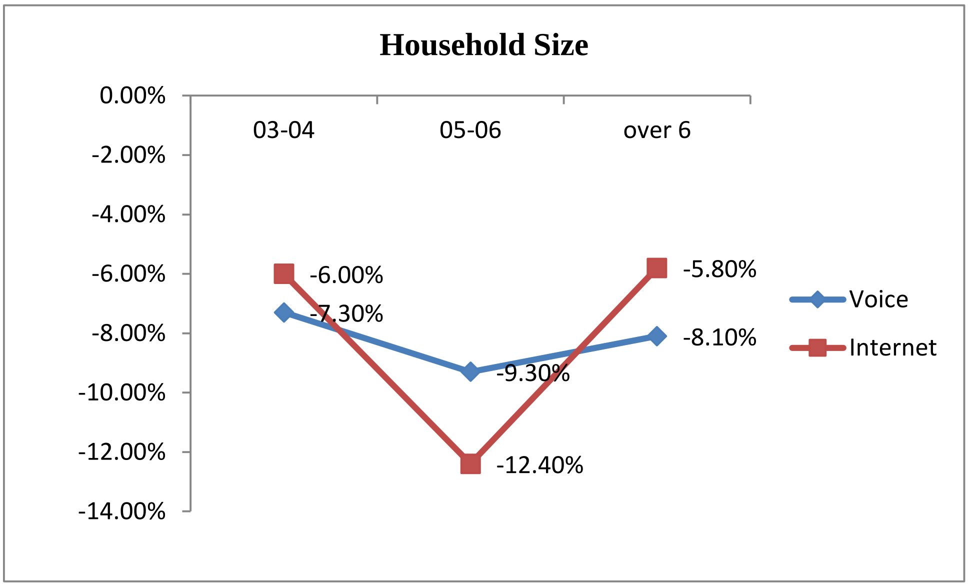


Figure 5.5 illustrates the risk of defection for customers subscribing to voice and Internet services. The horizontal axis shows the household size organised by the coded categories. The vertical axis indicates the defection rates. The negative coefficient indicates a lower defection rates.

Figure 5.5, for voice services shows that the risk of defection is 7.0 per cent lower for household size having 3 to 4 people, 9.0 per cent lower for household size having 5 to 6 people and 8 per cent lower for household size more than 6 people compared to the reference group. For Internet subscribers, the risk of defection is 6 per cent lower for household size having 3 to 4 people, 12.0 per cent lower for household size having 5 to 6 people and 6.0 per cent lower for household size more than 6 people compared to the reference group.

The findings indicate that the risk of defection is slightly lower for household size of 3 to 4 people subscribing to Internet services (6.0 per cent) than voice (7.0 per cent). However, the risk of defection is substantively lower for Internet services (12.0 per cent) than for voice (9.0 per cent) for household size of 5 to 6 people. For household size more than 6 people, the hazard for defection is lower in the Internet product (6.0 per cent) than the voice product. The findings could be interpreted as simply the effect of a difference between Internet and voice products and between the smaller and larger household size.

The possible explanation is that, unlike voice services, Internet access provides access to multiple members of the household at any point of time as compared to the generally single user at a time for voice product. Further, more users within a single household could make the decision to defect much more complex. Conversely, customers with smaller household sizes are more likely to defect and this can best be explained by the fact that larger households are more stable and may be less likely to move house and cancel their services. They are also more likely to be the targets of the firm's retention efforts because they tend to generate larger and more stable revenues for the firm. On the other hand, smaller households may change or cancel services due to job change, unemployment or move. These findings are similar to those of Vakratsas (1998) and (Buckinx & Van den Poel, 2005) who found that smaller households are more deal prone than larger sized households.

5.5 Examination of Interaction Effects

The primary aim of the research is to understand factors that influence customer retention behaviour; specifically using customer complaint behaviour. In an attempt to further estimate the retention behaviour, building upon the comprehensive model, Model 2, interaction effects among important predictors using backward selection were further evaluated in Model 3.

Interaction effects are used to examine the effect of one explanatory variable on the dependent variable where it depends on the magnitude of another independent variable (Ai & Norton, 2003). For example, the magnitude of customer survival or defection probability could be further estimated by examining the interaction between two explanatory variables such as effect of 'length of recovery' and 'severity of a problem'. Similarly the interactions of 'customer characteristic' explanatory variables could also refine our understanding of customer's retention by different segments. As an example, customer defection behaviour for a high income group and age may be different than those from a lower income and age.

Upon examination, all possible interaction terms were included using the commonly used stepwise backward elimination (Yuan & Lin, 2006). The selection technique, which is based on tests of statistical significance, helps to make reasonable decisions about which explanatory variables are most important and should be included in the model. The final explanatory variables included in the model are variables including: interaction between complainants and time since last complaint; length of recovery and severity; usage and age; and household income and age.

The results in Table 5.1 and Table 5.2 show that several explanatory variables had significant interaction terms. The rest were not included in the model, because the 'p value' was high and there was little improvement in model selection statistics AIC and SBC. The results also show that not all the explanatory variables are dependent on each other. For example, the parameter estimates for the interaction between usage and age for voice products is '0', indicating that the explanatory variable is independent of each other

Table 5.1 shows the estimated result for customer subscribing to voice services. For the interaction term between length of recovery*severity (Row 25), the result is significant at p<.0001 only for high severity cases. However, there was no support for interaction term length of recovery*low severity. The hazard ratio for interaction between length of recovery*severity high is 1.014. Table 5.2 shows the estimated result for customer subscribing Internet services. Similarly, the interaction term between length of recovery*severity (Row 25), the result is significant at p<.0001 only for high severity. No support was found for interaction term length of recovery*low severity. The hazard ratio for interaction between length of recovery*severity high is 1.011.

From the results it can be deduced that the risk of defection is higher 11.0 per cent and 14.0 per cent when the severity is categorised as high for voice and Internet services respectively. The results suggest that although customers experiencing high severity incidents have lower defection rates, these factors are independent of each other due to the service recovery paradox. Yet, prolonged recovery time from a systematic failure could reflect a cost to the customers and therefore increase the risk of defection. The findings may have managerial implications pointing as they do have key interaction between length of recovery and the severity of the incident on customer retention behaviour. Therefore, an action the firm might take is to speed up the recovery process, especially on high severity incidents because of the greater impact of high severity incidents versus low severity incidents.

Further examination of interaction on income*age and usage*age[39] were carried out for both the voice and Internet services. Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for interaction on income group medium*age group 2 (26-40 years), Row 19 is 1.196. The result is significant at p<.0001. The hazard ratio for interaction on income group medium*age group 2 (41-55 years), Row 20 is 1.062 but the results is not significant at p=0.1132. The hazard ratio for interaction on income group medium*age group 3 (>55 years), Row 21 is 0.937 and the results is significant at p=0.080.

Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for interaction on income group medium*age group 2 (26-40 years), Row 19 is 0.773. The hazard ratio for interaction on income group medium*age group 2 (41-55 years), Row 20 is 0.797. The results is significant at p<.0001. The hazard ratio for interaction on income group medium*age group 3 (>55 years), Row 21 is 0.896 and the results is significant at p=0.006.

[39] The interaction term usage*age for voice services is ignored since it is non-significant.

Figure 5.6: Interaction between Household Income (Medium) and Age

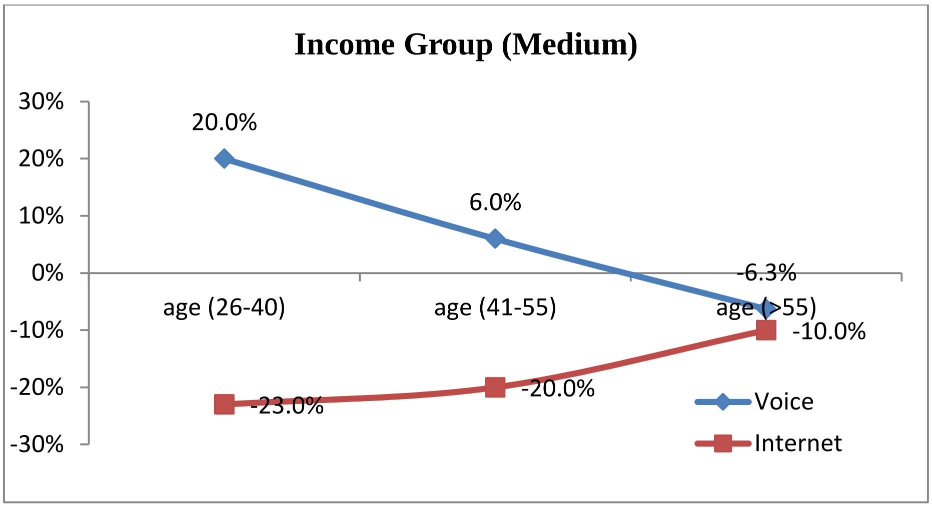


Figure 5.6 above illustrates the defection rates for interaction between medium income group and age across two product categories, voice and Internet. The horizontal axis is represented by the age group while the vertical axis provides the defection rates. The graph shows that the risk of defection for medium income customer subscribing to Internet services is 23.0 per cent lower for age group 2 (26-40) but the defection rates are 20.0 per cent higher when subscribing to voice services. Similar trend exist for customers subscribing to Internet services for age group 3 (41-55) the defection risk is 20.0 per cent lower but no inferences could be made for voice services because the p=0.1132. Finally for age group 4 (over 55) the risk of defection is lower for both customers subscribing to voice and Internet services.

Next, Table 5.1 shows the estimated result for customer subscribing voice services. The hazard ratio for interaction on income group high*age group 2 (26-40 years), Row 22 is 2.038. While, the hazard ratio for interaction on income group high*age group 3 (41-55 years), Row 23 is 1.802. Finally, the hazard ratio for interaction on income group high*age group 4 (over 55 years), 24 is 1.637. These results are significant at p<.0001.

From the results, it can be inferred that risk of defection is higher for households with high income and more pronounced in age group 2 (26-40 years).

Next, Table 5.2 shows the estimated result for customer subscribing Internet services. The hazard ratio for interaction on income group high*age group 2 (26-40 years), Row 22 is 1.278. While, hazard ratio for interaction on income group high*age group 3 (41-55 years), Row 23 is 1.408. Finally, the hazard ratio for interaction on income group high*age group 4 (over 55 years), Row 24 is 1.516. These results are significant at p<.0001.

Figure 5.7: Interaction between Household Income (High) and Age

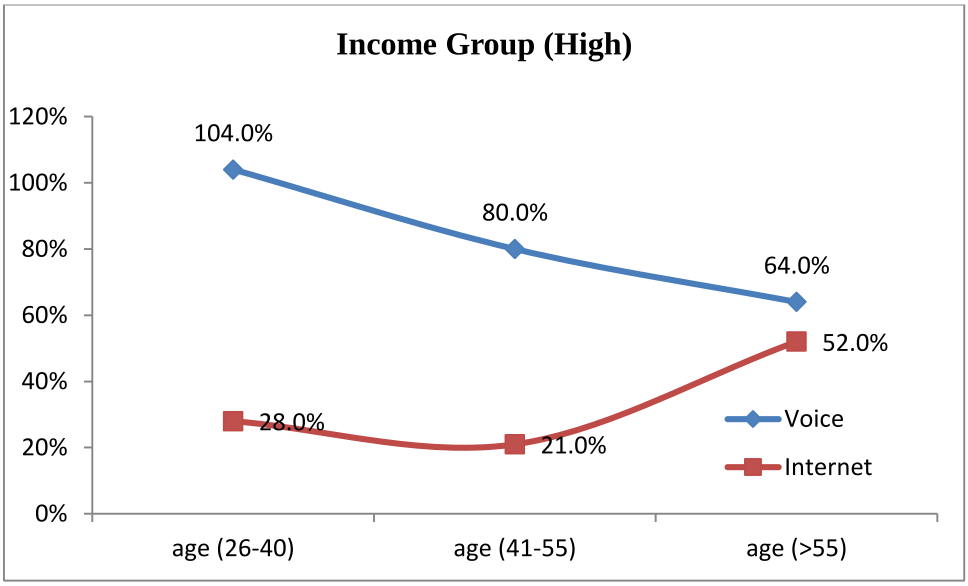


Figure 5.7 plots the defection rates for interaction between high income group and age across two product categories, voice and Internet. The horizontal axis is represented by the age group while the vertical axis provides the defection rates.

The graph shows that for Internet services the risk of defection is 28.0 per cent higher amongst customers who are between 26 and 40 years old. Whereas, for high income group, the risk of defection is 52.0 per cent greater amongst customers aged 55 years or more. Similar trends in risk of defecting also exist for the other age groups. For both the Internet and voice products, customers with higher household income are more likely to defect, and this is true even after adjusting household income by age.

Lastly, the interaction between the explanatory variable usage* age is significant for Internet but not for voice services. Table 5.2 shows the estimated result for customer subscribing to Internet services. The hazard ratio for interaction on usage*age group 2 (26-40 years), Row 26 is 0.999 and is significant at p=0.579. The hazard ratio for interaction on usage*age group 3 (41-55 years) is 0.999 and is significant at p=0.011. The hazard ratio for interaction on usage*age group 4 (over 55 years) is significant at p= 0.006. From the results, it can be inferred that risk of defection decreases on all category of age group by 1 per cent and this result suggest that high usage of the service reflect the users' approval of the service, as a result, these users are less likely to defect.

5.6 Summary

The models presented in this research capture the customer retention behaviour for two products at separate turning points in their life cycle. The models used for analysis are built on actual behaviour instead of the perceptual metrics used in previous research that have provided only a narrow understanding of retention behaviour over time. Clearly, the models provide significant insights into customer retention behaviour and make contributions to the existing literature on customer retention. Several explanatory variables are found to impact customer retention. When one or more of these explanatory variables change, the risk of the customer defection also changes. Interaction effects between selected variables are also tested, with some significant results.

To further refine understanding on customer retention behaviour, thirteen explanatory variables were included in the models hazard ratios). Each research question was examined against the results and among the more notable findings are the following:

- A customer who never complains is at greater risk of defection than one who has submitted a complaint and customer experiencing high severity has lower likelihood to defect.
- Across many categories, significant differences were found in the results for customers subscribing to voice services versus customers subscribing to Internet services, particularly for the demographic characteristic of household size.
- The interaction effects between selected variables, namely, recovery*severity, income*age, and, for Internet-service customers, usage*age, were found to be significant.

Chapter 6 further discusses the results, outlines theoretical and methodological contributions, and provides managerial and public policy implications. In addition, limitations of the present research are discussed, and suggestions for future research are made.

CHAPTER 6: DISCUSSION AND IMPLICATIONS

6.1 Introduction

This thesis examines the factors impacting customer retention and focuses on complaint behaviour for two products at different stages of the product life cycle. Few studies have focused on factors related to complaint behaviour. Furthermore, extant studies on retention have typically been carried out using perceptual metrics (i.e. self-reported behaviour), which provides a restrictive understanding of retention behaviour. The final chapter begins by summarising the purpose of this research. Next, theoretical implications of the results are discussed by considering factors that influence customer retention based on increased and decreased risk of defection. A comparison is then made of the two products studied in this research, voice services and Internet services, which are at different stages of the product life cycle. The effects of explanatory variables that increase or decrease the risk of defection are compared for both products. Next, methodological, managerial, and public policy implications are considered. Finally, limitations of the study are discussed and directions for future research are identified.

6.2 Aim of the study and research questions

This study uses a longitudinal and a semi-parametric approach to understand the effects that various explanatory variables may have on customer retention for two telecommunications products, voice and Internet services, which are at different stages in the product life cycle. A focus is made on complaint behaviour, which is measured at the individual level over multiple points in time. Previously identified drivers such as customer characteristics, usage, and relationship length were also included in the estimation model.

A semi-parametric approach was used to create the estimation model. To expand on previous research three Cox PH models were estimated for each product (i.e., voice and Internet). First, a baseline model (Model 1) was estimated using conventional drivers of customer retention. Second, a comprehensive model (Model 2) was estimated by including several time-variant explanatory variables. Third, an additional model was built based on the comprehensive model. As interactive models are important in longitudinal studies, even more so than cross-sectional ones because , this model was evaluated in greater detail by including interaction effects among the explanatory variables using backward selection and incorporating the significant interaction terms in to the comprehensive model (Model 3). Model 3 was found to have best fit (see Section 5.4.1 in Chapter 5) and is used in this study. Eight research questions, introduced in Chapter 1, are examined in detail in Chapter 5.

The results related to the key explanatory variables on expected causality are included in the comprehensive model with interactions (Model 3) are presented in Table 6.1. Key empirical findings related to the factors that impact customer retention behaviour are discussed in the following sections. The theoretical and managerial contributions are addressed, and the methodological advances offered in the estimation model are then outlined.

6.3 Theoretical implications

Table 6.1 is a summary of results for the explanatory variables in the comprehensive model, including interaction effects. The table presents the risk of defection for voice and Internet services. Overall, the results support customer complaint behaviour as a crucial dimension of the customer-firm relationship and customer retention. The key findings, organized by explanatory variable, are outlined in the following sections. Sections 6.3.1 and 6.3.2 discuss factors that influence customer retention behaviour toward increased and decreased risk of defection, respectively. Following this, Section 6.3.3 discusses the effects of the explanatory variables on risk of defection for voice services versus Internet services.

Table 6.1: Summary of key explanatory variables and interactions in the comprehensive model

Explanatory variable	Expected effect on customer retention based on extant literature	Finding
Complainant or non-complainants	Positive	Positive
Number of complaints	Positive or negative	Positive
Time since last complaint	Positive	Positive
Severity	Positive or negative	Positive
Length of recovery	Positive	Positive
Communication encounter	Positive	Negative
Length of relationship	Positive	Positive
Usage	Positive	Positive
Age	Positive	Negative
Gender	Positive or negative	Positive
Household income	Positive	Negative
Household size	Positive	Positive

6.3.1 Factors that influence customer retention based on increased risk of defection

6.3.1.1 Complainants vs. non-complainants

Extant literature has indicated that not all unhappy customers report their complaints (Chebat et al., 2005; Sharma et al., 2010). The risk of defection is increased when complaint behaviour is not reported and decreased when it is reported (Van den Poel & Larivière, 2004). This point was substantiated by US-based consumer and marketing consulting firm TARP, which stated that 54,0 per cent of the customers never complained when dealing service issues (Lovelock et al., 2011).

This research shows that non-complainants are at a greater risk of defecting than complainants. The risk of customer defection is 275.0 per cent for customers subscribing to Internet services, and 175.0 per cent for voice services. This finding parallels the finding by Lariviere and Van den Poel (2005) that customers who lodge complaints are, overall, “more active” or “more involved” with a firm and thus less likely to defect. Further, customers who complain may feel their complaints have been acknowledged by the firm and is likely to be acted upon. By this logic, customers who do not complain could have also experienced dissatisfaction but choose to express their complaints in other ways (e.g., by defecting to another firm or service or engaging in negative word of mouth behaviour). The findings are consistent with research that shows 50 to 80 per cent of non-complainants usually forego the opportunity to lodge complaints (Bogomolova, 2011; Goodwin & Ross, 1989).

6.3.1.2 Length of recovery

The work of Davidow (2003) supported timeliness of service recovery as one of the key factors in improving customer retention. A number of studies have operationalised the recovery effort within the context of customer satisfaction and used customer surveys to gauge customers' intention to continue with their service provider (Kau & Loh, 2006; Smith, Bolton, & Wagner, 1999b). Further, recovery time has been identified as an important criterion of procedural justice in complaint satisfaction that results in increased customer satisfaction and reduced churn (Blodgett et al., 1997; Clemmer & Schneider, 1996; Tax & Brown, 1998).

Findings from the present research indicate that the effects of the length of recovery concur with findings in previous studies that speed of recovery effort can be ascribed to customer defection behaviour. In particular, the study showed that the risk of defection increases by 4.0 per cent and 2.8 per cent for each extra day a firm fails to recover voice and internet services, respectively. That is, the higher the number of recovery days, the higher the propensity for a customer to defect. Although this finding agrees with those of Knox and Van Oest (2014) and Lariviere and Van den Poel (2005), the present research departs from some previous findings in that the length of recovery is operationalised using actual behaviour.

An additional dimension of this research is the evaluation of the interaction between length of recovery and incident severity on customer survival probability. The results suggest that the risk of defection is higher for internet and for voice services (1.1 per cent and 1.4 per cent, respectively) when severity is categorised as high; however, no effect was observed when severity of an incident was low. These findings indicate that although customers who have experienced high-severity incidents have lower defection rates, due to the service recovery paradox, prolonged recovery time from a system failure could be perceived as a cost to customers and hence increase their likelihood of defection

(Maxham & Netemeyer, 2002). The results show that timeliness of recovery is essential in environments such as the service sector and is especially acute in services such as telecommunications where reliability and availability are essential.

6.3.1.3 Communication encounter

In the literature, the communication encounter is referred to as frequency of contact or number of interactions per period between exchange partners (Doney & Cannon, 1997; Palmatier, Dant, Grewal, & Evans, 2006). This variable has been used to reflect relationship depth (Bolton, Lemon, & Verhoef, 2002), frequency of interaction (Homburg & Stock, 2004), and relationship experience (Bolton et al., 2000). So far, previous research supports the idea that more contact between a firm and its customers leads to higher retention. However, the findings of the present study suggest that the likelihood of defection increases by 2.2 per cent and 4.3 per cent for voice and Internet services, respectively, for every instance of communication with the firm.

The more pronounced effect for voice service than for Internet services may be because customer complaints (e.g., service failures, voice quality) have far greater impact than communication encounters for incidents such as inquiries, feedback, and customer requests. The results are best explained by the findings of Bitner (1990) that some key interactions between the customer and the firm can have relatively greater impact than other actions viewed as being of lesser impact. The results could also reflect that the firm's communication efforts and responsiveness to enquiries play a crucial role in shaping customer-firm encounters.

6.3.1.4 Customer demographics

Age

East et al. (2006) and Wright and Sparks (1999) revealed that older respondents tend to be less loyal than young or middle-aged customers. Conversely, Baumann et al. (2005) and Patterson (2007) found that older respondents tend to stay longer with their suppliers. Socio-emotional selective theory argues that older customers tend to have limited time and cognitive resources (Cole et al., 2008) and are thus less likely to develop new relationships with an alternate service provider.

For Internet services, middle-aged customers (26-40-year-olds) are at a 47.0 per cent risk of defecting compared to young customers (19-25-year-olds). Older customers (41-55-year-olds) have an even higher risk of defecting (73.0 per cent) compared with 19-25-year-olds. Furthermore, customers who are over 55 years of age are 70.0 per cent more likely to defect than19-25-year-olds. For voice services, the risk of defecting for middle-aged customers (26-40-year-olds) is 2.7 per cent higher than for young customers (19- to 25-year-olds). Although this is much smaller than the 47.0 per cent difference observed for internet services, it is still significant. Older customers still have a higher risk of defecting compared with young customers; in particular, 41- to 55-year-olds are 33.0 per cent more likely to defect than 19- to 25-year-olds, and customers who are over 55 are 41.0 per cent more likely to defect. The differences are less dramatic compared with similar figures for internet service. The findings from this research mirror those of East et al. (2006) and Wright and Sparks (1999).

The findings indicate that the risk of defection based on age is significantly different for voice and Internet services. For all age groups, risk of defection is much greater for

Internet services than for voice services. Because this is consistent across age groups, this indicates that other variables may be responsible for this difference.

The increased defection risk for middle-aged and older customers could be due to a number of possible reasons. As Internet and voice services are important to different age groups, these results differ according to the type of service (i.e., voice and Internet). First, the service packages offered by the firm in this study could be specifically designed for young people (who are more oriented toward Internet service). The firm may have already been aware that younger customers tend to churn and could have taken careful steps to make their younger customers happy and, in turn, retain them. Older customers not accustomed to new ways of doing business may be more likely to defect. Furthermore, the firm is likely aware that younger users are more oriented toward Internet services and may have marketed their packages as such. Second, it may be easier for young customers to access retention efforts such as online or email promotional offers, coupons and discounts as they are more technology savvy than older customers. Third, for people over 55 in particular, mortality could also be a reason for higher defection rates. Fourth, older customers may not be making decisions for themselves; instead, the decisions could be made by their children. Finally, a higher defection rate for middle-aged customers might be explained by lifestyle changes, for example, getting married, having children, buying a larger house, or moving to a different city, that occur more frequently when customers are within this age group. Customers experiencing these events may have changed needs and require different service plans to suit their new needs. Investigation of these reasons would require additional collection of data and is a topic for future research.

Gender

Melnyk et al. (2009) and Mittal and Kamakura (2001) found that female customers tend to be more loyal and are more risk averse than male customers. In contrast, Wirtz et al. (2014) found that females aged 50 years and older are more likely to switch. However, other studies by Jamal and Bucklin (2006) and Patterson (2007) found no significant correlation between gender and customer defection or loyalty.

The findings of the present research revealed that female customers are less likely to defect than male customers and that this is true for both internet and voice services. For voice services, females are 0.6 per cent less likely to defect than males. This result is not significant. For Internet services, females are 11.0 per cent less likely to defect than males. This difference may be due to minute details in firm operation such as marketing efforts and advertising's effects across gender groups. It is impossible to pinpoint the exact causes without considering more data such as the firm's marketing strategies and the occupations of its customers.

Household income

Income level has been found to affect customer decision-making and defection. Research has suggested that a high-income level is negatively correlated with customer defection (Keaveney & Parthasarathy, 2001; Reinartz & Kumar, 2003). The understanding is that high-income customers are less sensitive to price changes, especially in low-involvement purchases. Additionally, high-income customers may perceive higher switching costs as requiring too much time and effort.

However, the results from the present research are in contrast to those from existing research. They indicate that households with higher incomes tend to have a higher defection risk for both voice and internet services than medium- and low-income customers, even after adjusting household income by age. The results suggest that for high-income customers, the risk for defection is lower for voice services (19.0 per cent) than for internet services (62.0 per cent). For medium-income customers, the risk for defecting is also lower for voice services (27.0 per cent) than for internet services (46.0 per cent).

One possible explanation maybe that the firm in this study provides low-quality products/services at highly competitive prices, and high-income customers can afford and are willing to subscribe to better products/services at a higher price. As a result, high-income customers may be more likely to defect, and low-income customers may be more likely to stay with the firm. Moreover, people with higher incomes tend to be better educated and better at using the Web to search for competitive offers from other service providers, resulting in greater product involvement.

Household size

Previous studies have shown that large households have a lower likelihood of defecting than small households (Mittal & Kamakura, 2001; Risselada, Verhoef, & Bijmolt, 2010). The results of the present research suggest that for both voice and Internet services, customers in households with one or two people have a higher risk of defecting compared to customers in households with more than two people. A possible explanation for this finding is that unlike voice services, where access is limited to person-to-person connectivity, internet services provide access to multiple members of a household at any point of time. Furthermore, more users within a single household are involved in the decision to defect much more difficult. The higher likelihood of customers in smaller households to defect may also be explained by the fact that larger households are more stable and less likely to move around and are hence less likely to cancel their existing services. They are also more likely to be the target of a firm's retention efforts because they tend to generate large, stable revenue. On the other hand, smaller households may be more likely to change or cancel services due to job changes, unemployment or relocation. The research findings parallel those of Du, Kamakura, and Mela (2007) and Risselada et al., 2010 who found smaller households are more deal-prone than larger households.

6.3.2 Factors that influence customer retention based on decreased risk of defection

6.3.2.1 Number of complaints

The literature has shown that managing complaint volume is an important predictor of customer retention behaviour (Ahn et al., 2006; Bolton, 1998). As noted earlier, it is likely that number of complaints is associated with customer retention but the direction of the response parameter for number of complaints is uncertain. That is, the response parameter could be negatively or positively related to customer retention.

The research results of this study support the finding by Bolton (1998) that the risk of defection decreases as complaint volume increases. The total effect for the number of complaints is 5.0 per cent for voice and 15.0 per cent for internet services. The more complaints customers submit the more engaged they are with a product, making them less likely to defect if their complaints are addressed appropriately by the firm.

6.3.2.2 Recency of complaints

In the extant literature, 'recency' has been a focus of modelling within the context of renewal decisions and service experience (Bolton, Lemon and Bramlet, 2006). In the present research, the explanatory variable of time since last complaint was used as a proxy to measure the effect of recency. Results showed the risk of defection decreases by 2.4 per cent for internet and 3.9 per cent for voice services for every month that has elapsed since the previous complaint. Findings from past studies confirm that customers who complain about service failures within short time spans are likely to have lower satisfaction levels (Maxham & Netemeyer, 2002). Furthermore, other studies have shown that customers weigh recent incidents more heavily than overall outcomes in regard to their satisfaction (Ross & Simonson, 1991).

6.3.2.3 Severity of complaints

Studies have examined the effect of severity of complaints in regard to different outcomes as a result of context. For example, Maxham and Netemeyer (2002) examined failure severity and the impact on overall customer perception and repurchase intent. They found major, severe incidents influence satisfaction in a negative manner and decrease customer share.

Similarly, Smith and Bolton (1998) operationalised magnitude of severity as high and low in relation to hypothetical service failures and recovery encounters. They found that satisfaction and re-patronage intentions decreased for high severity cases compared to low severity cases. In contrast, Lariviere and Van den Poel (2005) focused on the relationship between post complaints and repurchase behaviour. Although they used "justification of complaints" as a surrogate to severity, they found that severe complaints (i.e., complaints with a higher level of justification) had a positive relationship with the complainant's next purchase decision. They attributed this to the consequence of recovery success and assumed more favourable recoveries for more justified service failures.

In the present research, the findings show that customers who had low-severity complaints were at greater risk of defection compared with non-complainants. The risk of defection is 375.0 per cent and 189.0 per cent for voice and internet services, respectively. These indicate that customers who complained about low-severity incidents were more likely to defect than customer who complained about high-severity incidents. However, customers with high-severity complaints were less likely to defect compared with non-complainants; the risk of defection is 58.0 per cent and 70.0 per cent lower. The research results contrast with those of Maxham and Netemeyer (2002) and Smith and Bolton (1998), who carried out research from the perspective of satisfaction and repurchase intention.

However, the results in this study mirror those of Lariviere and Van den Poel (2005) in terms of actual repurchase behaviour. The service recovery paradox can be used to explain the somewhat counterintuitive results; a higher level of favourable recoveries rather than failures indicates customers are more satisfied with their service after the resolution of a high-severity incident compared to how they would feel if the issue that prompted the complaint had never occurred The evaluation of service failure and recovery are important for customers subscribing to these utility-type services. The reliability and the functionality of the services can explain the differences between the two services.

6.3.2.4 Usage

Several previous studies have included the explanatory variable of usage in different contexts. For example, Bolton and Lemon (1999) studied usage patterns (customers' perception of fairness) and concluded that decreasing customer usage levels may be an indicator of eventual defection. Verhoef (2003) presented similar findings and showed that a high degree of usage, which may also be mirrored by share of wallet, is positively correlated to customers' decisions to continue using a company's products or services. Ahn et al. (2006) examined how service usage and customer status changes could be an early warning sign for potential defection. More recently, Ascarza and Hardie (2013) modelled usage within the context of contractual relationships, with a focus on predicting accuracy of usage and renewal behaviour.

The present research found that high usage levels have a positive effect on retention for Internet customers but a negative effect on retention for voice customers. A possible reason for this may be that usage of Internet services is on the rise whereas usage of voice services is declining. The results imply that each additional download per month corresponded to a 0.4 per cent reduction in risk of defection for internet customers and a 2.0 per cent increase in risk of defection for voice customers. The divergence of defection risk between voice and Internet services could also be driven by service attributes and indication of the utility of the service to customers. In other words, Internet services are on the rise and perceived by many as an essential service (like electric or gas services) whereas voice services are perceived by many as non-essential.

6.3.2.5 Length of relationship

A number of studies have found that the duration of customer-firm relationships is one of the key factors with a positive impact on retention (Bolton, 1998; Schweidel et al., 2008; Wirtz et al., 2014). As expected, for both voice and internet services, each additional month of a past relationship corresponds to a 0.4 per cent and 0.2 per cent reduction in the risk of defection, respectively. The data included in the models are for customers with a non-contractual relationship who are free to move to another firm without paying any fees, which further supports this finding.

This finding is as expected and supported by previous studies. The result suggests overall customer approval of the service since customers continuing their relationship with a firm signifies a high level of underlying involvement and overall utility of the services provided. Likewise, a customer who is unsatisfied with the utility of the services provided or the service quality is more likely terminate their relationship at an earlier stage. These findings are important to this research since the data included in the models are for customers in non-contractual relationships who are free to move to another firm without paying any switching fees.

6.3.3 Effects of explanatory variables on risk of defection for voice versus Internet service

The results of this study raise the question—do different products at different life-cycle stages with different market growth characteristics differ in terms of the effects of customer complaint behaviour, customer transactions, and demographics on customer defection for two distinct product markets. This section discusses effects in both directions. The effect size (i.e., for relevant explanatory variables, how great is the risk of defection within certain categories for voice services compared to Internet services) is also important, especially managerially.

Previous studies have suggested that customer retention is critical for mature markets (Athanassopoulos, 2000; Hogan et al., 2003; Jones et al., 2000). Thus, this research examined two products in the telecommunication market—voice and Internet services. Both products are relatively mature, but voice service is experiencing a decline while Internet service is still a growing market segment. Internet service is a growing market and there may be different effect sizes for the antecedents of customer defection compared to those for the mature/declining voice market.

This research examines these two products, which are at different life-cycle stages to identify differences in the antecedents of customer defection. To the best of the author's knowledge, no previous research has been conducted with regard to retention behaviour for two different products at different life-cycle stages using the explanatory variables in this study. Functionally, voice services and internet services have different product characteristics, and one service (voice) is in a declining stage and the other (internet) is in a growth stage.

6.3.3.1 Effect Direction

Table 6.2 below shows the comparison of increase risk and decreased risk of defection for two products, voice and Internet.

Table 6.2: Comparison between two Products (voice and Internet services) at Different Life Cycle Stages and Risk of Increased or Decreased Defection

Explanatory variable	Voice services		Internet services	
	Increases risk	Decreases risk	Increases risk	Decreases risk
Complainants vs. non complainants	✓		✓	
Number of complaints		✓		✓
Time since last complaint		✓		✓
Severity of complaints		✓		✓
Length of recovery	✓		✓	
Communication encounter	✓		✓	
Usage	✓			✓
Length of relationship		✓		✓
Age	✓		✓	
Gender		✓		✓
Household income	✓		✓	
Household size		✓		✓

From the table above, except for the explanatory variable of usage, the expected direction of the parameter estimates for both voice and internet are similar. From a utility standpoint, providing continuous service is essential for both voice and internet services. In terms of magnitude, the results demonstrate that internet customers generally have lower defection rates compared to voice customers. As noted by Rappa (2004), growth of information and communication platforms has made internet usage a mainstream activity. The explanatory variable of 'usage' is the only variable that differs significantly across the categories of voice and Internet usage (see Table 6.2). This suggests that customers consider Internet usage as perhaps the key aspect of their service, but realise access is dependent on external variables and expect issues regardless of the provider. Customer dependency on internet services may impact the customer's decision to stay with the firm.

Furthermore, rise in demand for online services such as paying of utilities bills, remittance, purchasing of media content could be a major factor in the differences of 'usage' for voice and Internet subscribers. This result may be further explained in that growing number of proportions of the voice calls are being made online via Voice over Internet Protocol (VoIP) (Business Monitor International)[40]. The characteristic and the manner of which subscriber use the services is an important factor accounting for differences in the defection rates between the two products. This could most likely impact customer's perceived value of the service and indirectly influence the usage behaviour (Peng & Wang, 2006).

[40] The information is sourced from Business Monitor International (BMI) telecommunication subscribed by Firm A. The report is part of BMI's copyrighted Industry Report & Forecasts Series dated January 2013.

Apart from the differences in terms of service attributes, price related attributes could explain the difference in usage between voice and Internet services. As noted by Lambrecht and Skiera (2006), specific pricing strategy are among other factors affects usage. Discussion with Firm A revealed that tariff plans for Internet and voice services are different. For voice services, the pricing is based on traditional structure of which subscribers pay a monthly fee and per minute chargers which are location and distance dependent. For Internet services, subscribers pay a flat monthly rate but maximum download will be capped by the type of plan they subscribe. Studies have shown either tiered tariff plan or flat tariff plan has direct influence to customer usage and ultimately their defection behaviour (Burnham et al., 2003; Danaher, 2002). Although, pricing is not examined in this research but it may be the case therefore that these variations of usage between voice and Internet services.

6.3.3.2 Effect Size

The products examined are at different stages in the product life cycle (i.e., voice services are in the decline stage and Internet services are in the growth stage) and accordingly there are differences in the antecedents of customer defection. A number of the explanatory variables had results with significantly different effect sizes on defection risk for voice services compared to Internet services. All categories of the customer characteristics of age and household income had greater effects on risk of defection for Internet services compared to voice services. A greater effect size can be observed across all categories of these two explanatory variables. Likewise, the customer complaint variables of complainant versus non-complainant and low severity of complaints also had greater effects on Internet services than on voice services. These results indicate that the service recovery paradox may be stronger for Internet services, which are increasingly considered essential for performing daily life activities, than for voice services. In contrast, the interaction effect of high-income and age had a greater effect on defection risk for voice services for the 26-40 and 41-55 age groups.

These findings have important managerial implications. While previous studies have suggested that customer retention is critical for mature markets (Athanassopoulos, 2000; Hogan et al., 2003; Jones et al., 2000), these results indicate certain factors may be even more crucial for growth markets. While some of the difference in effect sizes can be explained by the different product life-cycle stages of the products, functionally, voice and Internet services have different product characteristics, which may account for some of the variance between the products. Comparing retention behaviour across these two stages not only assists with proactive management of products but also helps with appropriate allocation of resources, depending on product stage.

Table 6.3: Comparison of effect size for voice and Internet services

Explanatory variables	Voice Risk of Defection	Internet Risk of Defection
Customer Charecteristic		
Age		
Group 1 (19-25) (reference group)		
Group 2 (26-40)	2.7%	47.0%
Group 3 (41-55)	33.0%	73.0%
Group 4 (over 55)	41.0%	70.0%
Gender (Female)	-0.6%	-11.0%
Household income		
low (reference group)		
medium	27.0%	46.0%
high	19.0%	63.0%
Household size		
01-02 (reference group)		
03-04	-7.3%	-6.0%
05-06	-9.3%	-12.40%
over 6	-8.1%	-5.8%
Customer Transaction		
Usage	2.0%	-0.4%
Length of relationship	-0.4%	-0.2%
Communication encounter	4.3%	2.2%
Customer Complaints		
Non-Complainants	175.0%	287.0%
Number of complaints	-4.9%	-15.7%
Time since last complaints	-3.9%	-2.4%
Severity		
none (reference group)		
High	-58.0%	-70.0%
Low	189.0%	375.0%
Length of recovery	-6%	0.40%
Interactions effects		
Income group (medium)*		
age (26-40)	20.0%	-23.0%
age (41-55)	6.0%	-20.0%
age (>55)	-6.3%	-10.0%
Income group (high)*		
age (26-40)	104.0%	28.0%
age (41-55)	80.0%	21.0%
age (>55)	64.0%	52.0%
Length of recovery *		
severity (high)	1.4%	1.1%
Usage *		
age (26-40)	0	-0.1%
age (41-55)	0	-0.1%
age (>55)	0	0.0%

6.4 Contributions

The aim of this research is to move beyond current knowledge in the area of customer retention. With this aim in mind, the study examined several factors that influence customer retention, with a specific focus on actual customer complaint behaviour. Some of these drivers have been previously examined, but extant research has largely ignored using the use of actual complaint behaviour. This research also studies products at different life-cycle stages, adding to the contribution of this research. The following subsections consider contributions in the areas of theory and methodology and provide implications for managers and public policy.

6.4.1 Theoretical Contributions

The present research is among a limited number of existing studies that have incorporated actual customer complaint behaviour as opposed to employing perceptual metrics. While Knox and Van Oest (2014) included customer complaints, their work differs in regard to factors such as severity and number of complaints, time effects, and other customer complaining/non-complaining behaviours. The present research confirms that customer complaint behaviour is influenced by the aforementioned explanatory variables. The research captured a solid representation of complaint behaviour through a dataset consisting of a complete census of actual consumer behaviour, which increased understanding of the factors impeding retention via means of decreasing or increasing risk of defection. Hence, the research enriches the existing service quality literature and refines the understanding of retention through the use of actual behavioural metrics.

The present research uses different models to expand on previous work in the customer retention literature. To analyse customer retention, the baseline model was initially employed and then expanded into the comprehensive and interaction models using selected explanatory variables. Previous work in customer retention behaviour (see Ahn et al., 2006; Coussement & Van den Poel, 2008; Hadden, Tiwari, Roy, & Ruta, 2006; Hung, Yen, & Wang, 2006; Kim & Yoon, 2004; Van den Poel & Larivière, 2004) has generally examined only a few factors instead of using integrated, comprehensive models to comprehensively examine customer retention behaviour. Further, most research has used models with parametric and non-parametric (Kamakura et al., 2005), which alone are insufficient to address the multitude of time-variant and time-invariant variables that affect retention behaviour, and cross-sectional approaches, which are insufficient to measure change over periods of time.

The present research addresses this gap in theory by using semi-parametric, longitudinal models to examine the effects of a variety of customer complaint behaviours as well as demographic characteristics and relationship metrics on defection risk. The empirical validation within a single dynamic hazard model allows the use of time-variant variables from different categories to assess the incidence and timing of customers' actual behaviour.

The findings in this research reaffirm the inclusion of some of the explanatory variables previously used to examine customer retention behaviour, including usage, relationship length, and customer characteristics. The research examined factors previously explored in the customer complaint and customer relationship marketing literature and including a number of these explanatory variables in our models. Specifically, the research focused on the inclusion of actual customer complaint behaviour, rather than stated intentions, as a way to examine customer retention behaviour. While initial retention models were built around survey data (Wong, 2011), advances in information technology and data storage have allowed the construction of models based on longitudinal data from a firm's client records, thus creating the opportunity to use operational rather than perceptual metrics (Braun & Schweidel, 2011). As perceptual measures of retention may not be an accurate surrogate for actual defection behaviour (Mittal & Kamakura, 2001), the present research enriches the existing service quality literature and refines the understanding of retention through the use of actual behavioural metrics.

Further, this research contributed to the wider relationship and services marketing literature, specifically customer retention studies, through a refined understanding of usage, relationship length, and customer characteristics. However, the most important aspect of this research stems from the holistic approach to understanding customer retention behaviour using data from several databases, which differs from the approaches in previous studies. The dataset is based on a sample size of between 15,000 and 20,000 active customers randomly selected from the customer base of two different products offered by an incumbent telecommunications firm. Randomisation of the sample over 36 months allowed for customer heterogeneity across different geographic locations. Longitudinal data on actual behaviour from a three-year period was utilized, allowing the model to capture simultaneous effects of multiple explanatory variables and, due to the granularity of the data, permitted comparison of results for two different products (i.e., voice services and Internet services). Most prior studies on defection using customer complaint behaviour have analysed from short periods of time (i.e., three months for Ahn et al. (2006)) with relatively small sample sizes (i.e., 171 by McQuilken et al. (2001), 201 by Hadden et al. (2006), 270 by Sharma et al. (2010), and 60 by Edvardsson and Roos (2001)).

Another important contribution of this research is the use of data on customers with non-contractual relationships with the firm. With a few exceptions (Reinartz and Kumar, 2003; Fader, Hardie, & Shang, 2010), most prior studies have been conducted in the context of contractual customer-firm relationships (Ascarza & Hardie, 2013; Bonfrer et al., 2010; Dover & Murthi, 2006; Schweidel et al., 2008). In contractual settings, customers are less likely to defect to avoid breaking a contract and paying related fees, introducing an additional perhaps unaccounted for variable into models. The use of data on customers with non-contractual relationships is another important contribution.

The present research expands on previous work and compares the effects of the included explanatory variables on two distinct products (voice services and Internet services). The use of the comprehensive model with interaction effects and the inclusion of actual customer behaviour data allowed the comparison of these two different products. To the best of the author's knowledge, this approach has not been taken in previous studies. This approach contributes to the existing body of knowledge on customer retention as the differences in customer retention behaviour for distinct products have not been previously examined in a comprehensive manner. Perhaps the most interesting finding from this research is that for the customer characteristics of age and household income, the customer complaint behaviour of complainant versus non-complainant and severity of the incident for which a complaint is lodged, and the interaction effect between income and age, the results for voice services and Internet services differed dramatically.

6.4.2 Methodological Implications

Many studies on customer retention have focused on using non-parametric and parametric modelling approaches. Although results from the parametric approach are relatively easy to interpret, this approach is limited in terms of incorporating unique marketing data characteristics that are inter temporal (Kamakura et al., 2005). This research filled the gap in regard to modelling approaches applied in customer retention research through inclusion of explanatory variables that are time-variant and time-invariant. More specifically, the research focused on using actual customer complaints to understand customer retention behaviour. The methodology employed in this research adopts a more rigorous approach through enables the inclusion of a large number of explanatory variables that are not addressed in the existing literature (Donkers et al., 2003). The inclusion of a large number of explanatory variables helps minimise bias that could contribute to over- or underestimating the parameter coefficients.

6.4.3 Managerial Implications

The managerial contributions of this research include multiple opportunities for firms to better understand how customer complaints impact their customer base and overall profitability. As the long-term survival and success of a business depends on how effectively it handles customer complaint behaviour, the comprehensive analysis of complaint data can serve as the impetus to change current operations to reflect a rapidly changing marketplace. As a broad example, organisations that conduct rigorous customer complaint analysis may be better able to change current operations for alignment with a rapidly changing, competitive, and unpredictable marketplace. Social media, rapid introduction of new technologies and broad changes in the ways people communicate make the managerial contributions of the present research more significant than ever.

In terms of organisational strategy, particularly marketing strategy, one of the main contributions of the present research is enabling prediction of customer defection behaviour. This predictive ability provides an opportunity to maximise a firm's allocation of limited financial and other resources across their product range and thus formulate a strategy for a more effective balance between retention and acquisition. The model created in the study could be further developed into a bespoke customer complaint handling system that could allow management to focus on improving or innovating quality of service and other processes to boost customer satisfaction. Furthermore, a targeted communication campaign could be designed to increase usage of existing products or services, perhaps by targeting customers showing a higher propensity for defection.

This research also presents management with an opportunity to utilise the nuanced and deeper understanding of customer complaint behaviour gained from comprehensive analysis for managing customer retention behaviour over time. The important findings of the research enable a firm to develop a robust customer care and complaint management system to efficiently handle complaints. In service industries, the managerial approach is essential in determining the effectiveness of customer relationships. Strategic relationship management and effective processes form a strong backbone for business sustainability and allow companies to gain a competitive edge. The firm has the opportunity to turn the initially adverse complaint experience into a strengthened customer relationship and an overall improvement in firm loyalty and brand image.

The current model allows firms to prioritize targeting customers with higher risk of defection in regard to providing appropriate products and services. A more complete understanding of customer retention behaviour provides an opportunity for firms to develop approaches for dealing with the loss of customers and potentially improve methods of retaining customers in the future. As firms have limited resources, this research can indirectly assist them in balancing their financial resources for acquisition and retention strategies. For example, firms may develop targeted and proactive communication campaigns to address customers who demonstrate a high propensity for defection. They could develop specific programs that increase usage of existing products or services.

Customers who are unsatisfied with a firm's service yet do not complain are a prime example of this. The results of the research allow for the prioritisation of customers who are unsatisfied with a firm's service quality yet (for various reasons) do not complain. Although they do not complain, these unsatisfied customers tend to be the least loyal (TARP, 2001). Apart from the risk of defection and becoming inactive, the non-complainants could also potentially damage the firm more broadly by spreading negative word of mouth amongst family, friends and colleagues.

Another customer cohort requiring specific attention is the inactive customers because maintaining inactive customers is costly for the firm. This research presents factors that the firm must attend to if they are to be successful in encouraging inactive customers to voice their dissatisfaction. This includes solving customer complaint issues in timely manner and, over time, extending their service relationship (e.g. bundled service offerings), ultimately preventing the defection of these inactive customers.

With the availability of vast amounts of consumer data and the utilization of complex algorithms, the latest technology and methods of analysis enable firms to monitor all aspects of customer satisfaction and retention technology, enabling the company to better serve customer needs. Major global firms leading the way in this arena include Google, Amazon, and Microsoft, all of which are obsessed with better understanding customer behaviour, even going as far as predicting behaviour before customers are aware of issues that may cause them to be unsatisfied, complain, or defect. Predictive activity such as this is possible only with the availability of complex algorithms and vast amounts of consumer data in short periods of time. The work in the present research advances predictive modelling, thus contributing to company's monitoring and prediction activity in regard to customer behaviour. A proactive managerial approach takes yet another step forward to predict areas for improvement or innovation to provide the next level of satisfaction for clientele.

This research demonstrates that managing quality customer data, exploring customer behaviour, and profiling and segmenting customers are all beneficial in managing retention and overall profitability. Managerial contributions flowing from this research potentially extend to several other elements of a firm's architecture and resources. For example, improving the alignment of human resource management systems with adequate focus on predicting customer defection behaviour. This might include appropriately configured incentive structures to align with marketing and customer retention strategies, goals and performance indicators. Organisation structure, leadership, staffing and training might also require alignment with the renewed customer retention focus. However, because of the mixed findings and the unique characteristics of the studied products (i.e., life-cycle stage), care should be taken when extending these findings to other industries and for products in different stages of the product life cycle.

In summary, proactive managerial skills are closely linked with the advancement of technology. The ubiquitous use of social media, changing ways in which people communicate, and trend toward the use of data to analyse previous customer behaviour and use this for predictive modelling of future behaviour make the present research particularly salient. Management skills evolve along research findings and the results of modelling. This entanglement is a known secret for many Fortune 500 companies to excel in their business, and remain relevant to their customers only because they are holding the needed information. It is essential that firms have the right information at the right time and use it in the right way to inform both strategy and action.

6.4.4 Public Policy Implications

In our rapidly world, it could be argued that telecommunication products are now perceived as a utility much the same as say gas or electricity. Telecommunications is now viewed as an essential service industry, one in which consumers and regulators are demanding higher standards of quality of service and overall customer care. There are some important public policy implications for this research.

Changes in public policy dealing with the relationship between telecommunications service providers and customers are resulting in regular reviews around the world. The

Telecommunications Consumer Protection Code (TCP Code) deal with a range of customer is used and, in the context of this research, customer complaint handling requires special emphasis. Although a registered code is technically regarded as subordinate legislation; there is typically no proactive requirement that exists for service providers to comply with any codes.

To provide focus to this discussion, the Australian context is used by way of example. In many instances the telecommunications industry is not obliged to comply with a code until it is specifically requested to do so, e.g., by the Australian Communications and Media Authority (ACMA) through a formal 'Direction to Comply'. Even then, such directions generally require compliance with a particular aspect of the code rather than the code in its entirety. It is important for this research to recognise the likely continuing strengthening of consumer protection regulations in the telecommunications industry as it relates to best practise customer complaints management. Additionally, some argue the need for improved integration between general and industry specific regulations and greater enforcement action.

Another important factor for firms to consider, in the current consumer protection regime, is the apparent difficulties in making the TCP Code accessible and comprehensible to customers. Reports indicate that few consumers even know the code exists, and even fewer appear able to comprehend the code sufficiently to be in a position to assert their rights to challenge their service provider because of complex and technically intense TCP Code descriptions. Some claim existing codes appear to privilege telecommunications service providers' interests rather than consumer rights.

There is also public policy implications related to Australian Consumer Law that applies many rules with respect to how firms must trade fairly with their consumers. With Government and regulators clearly calling for improved telecommunication consumer outcomes, ACMA launched the 'Reconnecting the Customer Inquiry' into complaints-handling and customer service. The ACMA Draft Report (2011) includes, amongst a range of broader solutions, recommendations to ensure customer care in the telecommunications.

Solutions leading to better complaints management and mechanisms to enable accurate comparison of service provider performance need to be encouraged by regulators. This study provides an opportunity for enterprises to better understand what represents outstanding service and to lead the market in terms of best practice customer care, in an industry where the incidence of complaints seems to be on the rise. If such initiatives become widespread, this could reduce the burden on public policy regulation and control. For example, it is likely that telecommunications regulators in the future will utilise customer complaint information as a means of developing a customer charter to balance the interests of the consumer and the telecommunications industry. Indeed, regulators may encourage customers to complain and then monitor the level of complaints received from them as another form of broad and firm specific consumer monitoring and protection.

In sum, this research provides telecommunication firms the opportunity to distinguish themselves from their peers by taking positive steps to enhance their customer care in recognition of the legislative preference by policymakers for self-regulation.

6.5 Limitations and Directions for Future Research

For most studies there are usually limitations. By examining customer complaint behaviour and other previously used explanatory variables that impact customer retention, this research provides you insights and contributes to a greater understanding of customer retention behaviour. A longitudinal approach was employed in the research to understand customer retention behaviour, specifically using customer complaint behaviour. Although this research helps explain the factors that influence retention behaviour, combining attitudinal inputs obtained through customer surveys would permit additional insights. However, collecting this type of information could be costly and would involve a lengthy process.

This research focused on examining the customer retention behaviour of consumers using two products, voice services and Internet services, at two different stages of the product life cycle (i.e., the growth and decline life-cycle stages). It would be interesting to extend this research to study services that are in other stages of the life-cycle, e.g., in the introduction and more mature stages of the product lifecycle.

Much of the research on customer retention (e.g. Thompson & Sinha, 2008; Ahn et al. (2006) has generally modelled growing markets. Scant attention has been given to modelling customer retention behaviour in declining markets, which is addressed in this study. However, comparing customer retention behaviour for product across the four stages of the product life cycle stages (i.e., introduction, growth, maturity and decline) may provide additional insights into differences and similarities in the effect direction and effect size of the key variables that impact customer retention.

Furthermore, the estimation approach employed in this research was to examine factors that impede customer retention using customer complaint behaviour and other behavioural measures. This approach was adopted to explain how these explanatory variables impede customer retention vis á vis decreased or increased risk for defection. A natural extension of the research would explore predictive models. However, a challenge in creating an accurate predictive model is whether a model estimated at a given time can predict customer retention at a later point in time. An additional challenge is determining which methodological approaches have the highest probability of accuracy and success.

Finally, the research focused on two subscription-based services, namely for telecommunications services. Due to the inherently different nature of the two products, the generalizing of the findings and related implications to other industries should be done with care. The research could be expanded to other industries such as insurance and finance to determine if the factors identified in this research also hold true in other industries. This would permit greater generalisability of the findings.

REFERENCES

Ahmad, R., & Buttle, F. (2001). Customer retention: A potentially potent marketing management strategy. *Journal of Strategic Marketing, 9*(1), 29-45.

Ahn, J. H., Han, S. P., & Lee, Y. S. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications Policy, 30*(10), 552-568.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters, 80*(1), 123-129.

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics, 16*(1), 3-14.

Allison, P. D. (2010). *Survival Analysis Using SAS: A Practical Guide* (2nd Edition). Cary, North Carolina: SAS Institute.

Alvarez, L. S., Casielles, R. V., & Martin, A. M. D. (2010). Analysis of the role of complaint management in the context of relationship marketing. *Journal of Marketing Management, 27*(1-2), 143-164.

Ascarza, E., & Hardie, B. G. (2013). A joint model of usage and churn in contractual settings. *Marketing Science, 32*(4), 570-590.

Athanassopoulos, A. D. (2000). Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of Business Research, 47*(3), 191-207.

Aydin, S., Özer, G., & Arasil, Ö. (2005). Customer loyalty and the effect of switching costs as a moderator variable: A case in the Turkish mobile phone market. *Marketing Intelligence & Planning, 23*(1), 89-103.

Bansal, H. S., & Taylor, S. F. (1999). The service provider switching model (SPSM): A model of consumer switching behavior in the services industry. *Journal of Service Research, 2*(2), 200-218.

Baumann, C., Burton, S., & Elliott, G. (2005). Determinants of customer loyalty and share of wallet in retail banking. *Journal of Financial Services Marketing, 9*(3), 231-248.

Berry, L. L. (1983). Relationship marketing. Emerging perspectives on service marketing (eds. L. L. Berry, G. L. Shostack, G. D. Upah), Chicago, IL: American Marketing Association, 25–48.

Berry, L. L. (1995). Relationship marketing of services—growing interest, emerging perspectives. *Journal of the Academy of Marketing Science, 23*(4), 236-245.

Berry, M. J., & Linoff, G. S. (2004). *Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management*. Indianapolis: John Wiley & Sons.

Bhote, K. R., & Bhote, A. K. (2004). *World class reliability: Using Multiple Environment Overstress Tests to Make it Happen.* New York: AMACOM.

Bitner, M. J. (1990). Evaluating service encounters: The effects of physical surroundings and employee responses. *Journal of Marketing*, *54*(2), 69-82.

Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). *Database Marketing: Analyzing and Managing Customers*. New York, NY: Springer.

Blodgett, J. G., & Anderson, R. D. (2000). A Bayesian network model of the consumer complaint process. *Journal of Service Research, 2*(4), 321-338.

Blodgett, J. G., Hill, D. J., & Tax, S. S. (1997). The effects of distributive, procedural, and interactional justice on postcomplaint behavior. *Journal of Retailing, 73*(2), 185-210.

Bogomolova, S. (2011). Service quality perceptions of solely loyal customers. *International Journal of Market Research, 53*(6), 793-810.

Bogomolova, S., & Romaniuk, J. (2009). Brand defection in a business-to-business financial service. *Journal of Business Research, 62*(3), 291-296.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with a continuous service provider: The role of satisfaction. *Marketing Science, 17*(1), 45-65.

Bolton, R. N., Kannan, P. K., & Bramlett, M. D. (2000). Implications of loyalty program membership and service experiences for customer retention and value. *Journal of the Academy of Marketing Science, 28*(1), 95-108.

Bolton, R. N., & Lemon, K. N. (1999). A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction. *Journal of Marketing Research, 36*(2), 171-186.

Bolton, R. N., Lemon, K. N., & Bramlett, M. D. (2006). The effect of service experiences over time on a supplier’s retention of business customers. *Management Science, 52*(12), 1811-1823.

Bolton, R. N., Lemon, K. N., & Verhoef, P. C. (2004). The theoretical underpinnings of customer asset management: A framework and propositions for future research. *Journal of the Academy of Marketing Science, 32*(3), 271-292.

Bonfrer, A., Knox, G., Eliashberg, J., & Chiang, J. (2010). A first-passage time model for predicting inactivity in a contractual setting. *Available at SSRN 997810*.

Boone, L. E., & Kurtz, D. L. (2011). *Contemporary Marketing* (15th ed.). Mason, Ohio: South-Western, Cengage Learning.

Borle, S., Singh, S. S., & Jain, D. C. (2008). Customer lifetime value measurement. *Management Science, 54*(1), 100-112.

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political analysis*, *14(*1), 63-82.

Braun, M., & Schweidel, D. A. (2011). Modeling customer lifetimes with multiple causes of churn. *Marketing Science, 30*(5), 881-902.

Brock, C., Blut, M., Evanschitzky, H., & Kenning, P. (2013). Satisfaction with complaint handling: A replication study on its determinants in a business-to-business context. *International Journal of Research in Marketing, 30*(3), 319-322.

Brockett, P. L., Golden, L. L., Guillen, M., Nielsen, J. P., Parner, J., & Perez-Marin, A. M. (2008). Survival analysis of a household portfolio of insurance policies: How much time do you have to stop total customer defection? *Journal of Risk and Insurance, 75*(3), 713-737.

Buckinx, W., & Van den Poel, D. (2005). Customer base analysis: Partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research, 164*(1), 252-268.

Burez, J., & Van den Poel, D. (2007). CRM at a pay-TV company: Using analytical models to reduce customer attrition by targeted marketing for subscription services. *Expert Systems with Applications, 32*(2), 277-288.

Burnham, T. A., Frels, J. K., & Mahajan, V. (2003). Consumer switching costs: A typology, antecedents, and consequences. *Journal of the Academy of Marketing Science, 31*(2), 109-126.

Chandrashekaran, M., Rotte, K., Tax, S. S., & Grewal, R. (2007). Satisfaction strength and customer loyalty. *Journal of Marketing Research, 44*(1), 153-163.

Chebat, J.-C., Davidow, M., & Codjovi, I. (2005). Silent voices why some dissatisfied consumers fail to complain. *Journal of Service Research, 7*(4), 328-342.

Choi, S., & Mattila, A. S. (2008). Perceived controllability and service expectations: Influences on customer reactions following service failure. *Journal of Business Research, 61*(1), 24-30.

Clemmer, E. C., & Schneider, B. (1996). Fair service. *Advances in Services Marketing and Management, 5*, 109-126.

Cole, C., Laurent, G., Drolet, A., Ebert, J., Gutchess, A., Lambert-Pandraud, R., Mullet, E., Norton,M., & Peters, E. (2008). Decision making and brand choice by older consumers. *Marketing Letters, 19*(3-4), 355-365.

Colgate, M., Tong, V. T.-U., Lee, C. K.-C., & Farley, J. U. (2007). Back from the brink: Why customers stay. *Journal of Service Research, 9*(3), 211-228.

Conlon, D. E., & Murray, N. M. (1996). Customer perceptions of corporate responses to product complaints: The role of explanations. *Academy of Management Journal, 39*(4), 1040-1056.

Coussement, K. (2014). Improving customer retention management through cost-sensitive learning. *European Journal of Marketing, 48*(3/4), 477-495.

Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. *Expert Systems with Applications, 34*(1), 313-327.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187-220.

Cox, D. R., & Oakes, D. (1984). *Analysis of Survival Data* (1st Edition). London: Chapman & Hall/CRC.

Cox, D.R., & Wermuth, N. (1996). *Multivariate Dependencies: Models, Analysis and Interpretation*. Boca Raton, Florida: Chapman and Hall/CRC

Cranage, D. (2004). Plan to do it right and plan for recovery. *International Journal of Contemporary Hospitality Management, 16*(4), 210-219.

Crosby, L. A., Evans, K. R., & Cowles, D. (1990). Relationship quality in services selling: An interpersonal influence perspective. *Journal of Marketing*, *54*(3), 68-81.

Crosby, L. A., & Stephens, N. (1987). Effects of relationship marketing on satisfaction, retention, and prices in the life insurance industry. *Journal of Marketing Research*, *24*(4), 404-411.

Dabholkar, P. A., & Walls, S. (1999). Service evaluation and switching behavior for experiential services: An empirical test of gender differences within a broader conceptual framework. *Journal of Consumer Satisfaction Dissatisfaction and Complaining Behavior, 12*, 123-137.

Danaher, P. J. (2002). Optimal pricing of new subscription services: Analysis of a market experiment. *Marketing Science, 21*(2), 119-138.

Davidow, M. (2003). Organizational responses to customer complaints: What works and what doesn't. *Journal of Service Research, 5*(3), 225-250.

De Matos, C. A., Henrique, J. L., & Rossi, C. A. V. (2007). Service recovery paradox: A meta-analysis. *Journal of Service Research, 10*(1), 60-77.

del Río-Lanza, A. B., Vázquez-Casielles, R., & Díaz-Martín, A. M. (2009). Satisfaction with service recovery: Perceived justice and emotional responses. *Journal of Business Research, 62*(8), 775-781.

Doney, P. M., & Cannon, J. P. (1997). An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing*, *61*(2), 35-51.

Donkers, B., Franses, P. H., & Verhoef, P. C. (2003). Selective sampling for binary choice models. *Journal of Marketing Research, 40*(4), 492-497.

Dover, H. F., & Murthi, B. (2006). Asymmetric effects of dynamic usage behavior on duration in subscription-based online service. *Journal of Interactive Marketing, 20*(3-4), 5-15.

Dowling, G. R., & Uncles, M. (1997). Do customer loyalty programs really work? *Sloan Management Review, 38*, 71-82.

Du, J., Fan, X., & Feng, T. (2014). Group emotional contagion and complaint intentions in group service failure: The role of group size and group familiarity. *Journal of Service Research, 17*(3), 326-338.

Du, R. Y., Kamakura, W. A., & Mela, C. F. (2007). Size and share of customer wallet. *Journal of Marketing, 71*(2), 94-113.

Dwyer, F. R. (1989). Customer lifetime valuation to support marketing decision making. *Journal of Direct Marketing, 3*(4), 8-15.

Dwyer, F. R., Schurr, P. H., & Oh, S. (1987). Developing buyer-seller relationships. *Journal of Marketing*, *51*(2), 11-27.

East, R., Hammond, K., & Gendall, P. (2006). Fact and fallacy in retention marketing. *Journal of Marketing Management, 22*(1-2), 5-23.

Edvardsson, B., & Roos, I. (2001). Critical incident techniques: Towards a framework for analysing the criticality of critical incidents. *International Journal of Service Industry Management, 12*(3), 251-268.

Eriksson, K., & Vaghult, A. L. (2000). Customer retention, purchasing behavior and relationship substance in professional services. *Industrial Marketing Management, 29*(4), 363-372.

Fader, P. S., & Hardie, B. G. (2010). Customer-base valuation in a contractual setting: The perils of ignoring heterogeneity. *Marketing Science, 29*(1), 85-93.

Fader, P. S., Hardie, B. G., & Lee, K. L. (2005). RFM and CLV: Using so-value curves for customer base analysis. *Journal of Marketing Research, 42*(4), 415-430.

Fader, P. S., Hardie, B. G., & Shang, J. (2010). Customer-base analysis in a discrete-time noncontractual setting. *Marketing Science, 29*(6), 1086-1108.

Fang, Z., Luo, X., & Jiang, M. (2013). Quantifying the dynamic effects of service recovery on customer satisfaction evidence from Chinese mobile phone markets. *Journal of Service Research, 16*(3), 341-355.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: nature, purpose, and findings. *Journal of Marketing*, *60*(4), 7-18.

Fornell, C., & Wernerfelt, B. (1987). Defensive marketing strategy by customer complaint management: A theoretical analysis. *Journal of Marketing Research*, *24(*4), 337-346.

Fornell, C., & Wernerfelt, B. (1988). A model for customer complaint management. *Marketing Science, 7*(3), 287-298.

Gelbrich, K., Gäthke, J., & Grégoire, Y. (2015). How much compensation should a firm offer for a flawed service? An examination of the nonlinear effects of compensation on satisfaction. *Journal of Service Research, 18*(1), 107-123.

Gelman, A., & Stern, H. (2006). The difference between "significant" and "not significant" is not itself statistically significant. *The American Statistician, 60*(4), 328-331.

Gilula, Z., McCulloch, R. E., & Rossi, P. E. (2006). A direct approach to data fusion. *Journal of Marketing Research, 43*(1), 73-83.

Giudicati, G., Riccaboni, M., & Romiti, A. (2013). Experience, socialization and customer retention: Lessons from the dance floor. *Marketing Letters,24*(4), 409-422.

Golder, P. N., & Tellis, G. J. (2004). Growing, growing, gone: Cascades, diffusion, and turning points in the product life cycle. *Marketing Science, 23*(2), 207-218.

Goodwin, C., & Ross, I. (1989). Salient dimensions of perceived fairness in resolutions of service complaints. *Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior, 2*, 87-92.

Grambsch, P. M., & Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*(3), 515-526.

Grönroos, C. (1994). Quo vadis, marketing? Toward a relationship marketing paradigm. *Journal of Marketing Management, 10*(5), 347-360.

Grönroos, C. (1990). Relationship approach to marketing in service contexts: The marketing and organizational behavioral interface. *Journal of Business Research, 20*(1), 3-11.

Gummesson, E. (1994). Making relationship marketing operational. *International Journal of Service Industry Management, 5*(5), 5-20.

Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., & Sriram, S. (2006). Modeling customer lifetime value. *Journal of Service Research, 9*(2), 139-155.

Gupta, S., & Lehmann, D. R. (2003). Customers as assets. *Journal of Interactive Marketing, 17*(1), 9-24.

Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. *Journal of Marketing Research, 41*(1), 7-18.

Gupta, S., & Zeithaml, V. (2006). Customer metrics and their impact on financial performance. *Marketing Science, 25*(6), 718-739.

Hadden, J., Tiwari, A., Roy, R., & Ruta, D. (2006). *Churn Prediction Using Complaints Data.* Paper presented at the Proceedings of World Academy of Science, Engineering and Technology.

Haj-Salem, N., & Chebat, J.-C. (2014). The double-edged sword: The positive and negative effects of switching costs on customer exit and revenge. *Journal of Business Research, 67*(6), 1106-1113.

Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: An empirical study. *International Journal of Service Industry Management, 7*(4), 27-42.

Hammond, K. (1998). Customer defection from supermarkets. *Advances in Consumer Research, 25*, 507-512.

Harker, M. J. (1999). Relationship marketing defined? An examination of current relationship marketing definitions. *Marketing Intelligence & Planning, 17*(1), 13-20.

Hart, C. W., Heskett, J. L., & Sasser Jr, W. E. (1990). The profitable art of service recovery. *Harvard Business Review, 68*(4), 148.

Helsen, K., & Schmittlein, D. C. (1993). Analyzing duration times in marketing: Evidence for the effectiveness of hazard rate models. *Marketing Science, 12*(4), 395-414.

Hertz-Picciotto, I., & Rockhill, B. (1997). Validity and efficiency of approximation methods for tied survival times in Cox regression. *Biometrics, 53*(3), 1151-1156.

Heskett, J. L., & Schlesinger, L. (1994). Putting the service-profit chain to work. *Harvard Business Review, 72*(2), 164-174.

Heung, V. C., & Lam, T. (2003). Customer complaint behaviour towards hotel restaurant services. *International Journal of Contemporary Hospitality Management, 15*(5), 283-289.

Hogan, J. E., Lemon, K. N., & Libai, B. (2003). What is the true value of a lost customer? *Journal of Service Research, 5*(3), 196-208.

Homburg, C., & Fürst, A. (2005). How organizational complaint handling drives customer loyalty: An analysis of the mechanistic and the organic approach. *Journal of Marketing, 69*(3), 95-114.

Homburg, C., & Stock, R. M. (2004). The link between salespeople's job satisfaction and customer satisfaction in a business-to-business context: A dyadic analysis. *Journal of the Academy of Marketing Science, 32*(2), 144-158.

Hung, S.-Y., Yen, D. C., & Wang, H.-Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications, 31*(3), 515-524.

Jain, D., & Singh, S. S. (2002). Customer lifetime value research in marketing: A review and future directions. *Journal of Interactive Marketing, 16*(2), 34-46.

Jain, D. C., & Vilcassim, N. J. (1991). Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Science, 10*(1), 1-23.

Jamal, Z., & Bucklin, R. E. (2006). Improving the diagnosis and prediction of customer churn: A heterogeneous hazard modeling approach. *Journal of Interactive Marketing, 20*(3-4), 16-29.

Jenkins, S. P. (1995). Easy estimation methods for discrete-time duration models. *Oxford Bulletin of Economics and Statistics, 57*(1), 129-136.

Johnston, R. (2001). Linking complaint management to profit. *International Journal of Service Industry Management, 12*(1), 60-69.

Joireman, J., Grégoire, Y., Devezer, B., & Tripp, T. M. (2013). When do customers offer firms a "second chance" following a double deviation? The impact of inferred firm motives on customer revenge and reconciliation. *Journal of Retailing, 89*(3), 315-337.

Jones, M. A., Mothersbaugh, D. L., & Beatty, S. E. (2000). Switching barriers and repurchase intentions in services. *Journal of Retailing, 76*(2), 259-274.

Kalamas, M., Laroche, M., & Makdessian, L. (2008). Reaching the boiling point: Consumers' negative affective reactions to firm-attributed service failures. *Journal of Business Research, 61*(8), 813-824.

Kaltcheva, V. D., Winsor, R. D., & Parasuraman, A. (2013). Do customer relationships mitigate or amplify failure responses? *Journal of Business Research, 66*(4), 525-532.

Kamakura, W., Mela, C. F., Ansari, A., Bodapati, A., Fader, P., Iyengar, R., Naik, P., Neslin, S., Sun, B., Verhoef, P.C., Wedel, M., & Wilcox, R. (2005). Choice models and customer relationship management. *Marketing Letters, 16*(3-4), 279-291.

Kamakura, W. A., Wedel, M., De Rosa, F., & Mazzon, J. A. (2003). Cross-selling through database marketing: A mixed data factor analyzer for data augmentation and prediction. *International Journal of Research in Marketing, 20*(1), 45-65.

Kau, A.-K., & Loh, W.-Y. E. (2006). The effects of service recovery on consumer satisfaction: A comparison between complainants and non-complainants. *Journal of Services Marketing, 20*(2), 101-111.

Keaveney, S. M. (1995). Customer switching behavior in service industries: An exploratory study. *Journal of Marketing, 59*(2), 71-82.

Keaveney, S. M., & Parthasarathy, M. (2001). Customer switching behavior in online services: An exploratory study of the role of selected attitudinal, behavioral, and demographic factors. *Journal of the Academy of Marketing Science, 29*(4), 374-390.

Keep, W. W., & Schneider, G. P. (2010). Deception and defection from ethical norms in market relationships: a general analytic framework. *Business Ethics: A European Review, 19*(1), 64-80.

Keiningham, T. L., Morgeson, F. V., Aksoy, L., & Williams, L. (2014). Service failure severity, customer satisfaction, and market share: An examination of the airline industry. *Journal of Service Research, 17*(4), 415-431.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal Relations: A Theory of Interdependence.* New York, NY: John Wiley and Sons.

Kelley, S. W., Hoffman, K. D., & Davis, M. A. (1994). A typology of retail failures and recoveries. *Journal of Retailing, 69*(4), 429-452.

Kim, H.-S. (2000). The churn analysis and determinants of customer loyalty in Korean mobile phone. *Korean Information Society Review, 2000,* 1-18.

Kim, H.-S., & Yoon, C.-H. (2004). Determinants of subscriber churn and customer loyalty in the Korean mobile telephony market. *Telecommunications Policy, 28*(9), 751-765.

Knox, G., & Van Oest, R. (2014). Customer complaints and recovery effectiveness: A customer base approach. *Journal of Marketing, 78*(5), 42-57.

Kumar, V., Bohling, T. R., & Ladda, R. N. (2003). Antecedents and consequences of relationship intention: Implications for transaction and relationship marketing. *Industrial Marketing Management, 32*(8), 667-676.

Kumar, V., & Luo, A. M. (2008). Integrating purchase timing, choice, and quantity decisions models. *Review of Marketing Research*, *28*(4), 63-91.

Lambrecht, A. & Skiera, B. 2006. Paying too much and being happy about it: Existence, causes and consequences of tariff-choice biases. *Journal of Marketing Research,* 43**,** 212-223.

Landon Jr, E. L. (1980). The direction of consumer complaint research. *Advances in Consumer Research, 7*(1), 335-338.

Lariviere, B., & Van den Poel, D. (2005). Investigating the post-complaint period by means of survival analysis. *Expert Systems with Applications, 29*(3), 667-677.

Laufer, D. (2002). Are antecedents of consumer dissatisfaction and consumer attributions for product failures universal. *Advances in Consumer Research, 29*(1), 312-317.

Lawless, J. F. (2011). *Statistical Models and Methods for Lifetime Data* (2nd Edition). Hoboken, NJ: John Wiley & Sons.

Lee, J., Lee, J., & Feick, L. (2001). The impact of switching costs on the customer satisfaction-loyalty link: Mobile phone service in France. *Journal of Services Marketing, 15*(1), 35-48.

Lemke, F., Clark, M., & Wilson, H. (2011). Customer experience quality: An exploration in business and consumer contexts using repertory grid technique. *Journal of the Academy of Marketing Science, 39*(6), 846-869.

Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, *43*(2), 276-286.

Lewis, B. R., & McCann, P. (2004). Service failure and recovery: Evidence from the hotel industry. *International Journal of Contemporary Hospitality Management, 16*(1), 6-17.

Liu, R. R., & McClure, P. (2001). Recognizing cross-cultural differences in consumer complaint behavior and intentions: An empirical examination. *Journal of Consumer Marketing, 18*(1), 54-75.

Lovelock, C. H. (1983). Classifying services to gain strategic marketing insights. *Journal of Marketing*, *47*(3), 9-20.

Lovelock, C. L., Walker, R. H., & Patterson, P. G. (2011). *Services Marketing: An Asia-Pacific Perspective* (5th ed.). Sydney: Pearson Education Australia.

Malhotra, A., & Kubowicz Malhotra, C. (2013). Exploring switching behavior of US mobile service customers. *Journal of Services Marketing, 27*(1), 13-24.

Marshall, G. W., & Shepherd, C. D. (1999). Selling and sales management in action: Service quality and the sales force: A tool for competitive advantage. *Journal of Personal Selling & Sales Management, 19*(3), 73-82.

Maxham III, J. G. (2001). Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. *Journal of Business Research, 54*(1), 11-24.

Maxham III, J. G., & Netemeyer, R. G. (2002). A longitudinal study of complaining customers' evaluations of multiple service failures and recovery efforts. *Journal of Marketing, 66*(4), 57-71.

Maxham III, J. G., & Netemeyer, R. G. (2002). Modeling customer perceptions of complaint handling over time: The effects of perceived justice on satisfaction and intent. *Journal of Retailing, 78*(4), 239-252.

McQuilken, L. (2010). The influence of failure severity and employee effort on service recovery in a service guarantee context. *Australasian Marketing Journal, 18*(4), 214-221.

McQuilken, L., Breth, R., & Shaw, R. (2001). *Relationship marketing and pay TV: satisfaction, complaining behavior, service recovery, and repurchase.* Paper presented at the Proceedings of the 9th International Colloquium in Relationship Marketing.

Melnyk, V., Van Osselaer, S. M. J., & Bijmolt, T. H. A. (2009). Are women more loyal customers than men? Gender differences in loyalty to firms and individual service providers. *Journal of Marketing, 73*(4), 82-96.

Miller, J. L., Craighead, C. W., & Karwan, K. R. (2000). Service recovery: A framework and empirical investigation. *Journal of Operations Management, 18*(4), 387-400.

Mills, M. (2011). *Introducing Survival and Event History Analysis. London, UK: Sag*e Publications.

Mittal, V., Huppertz, J. W., & Khare, A. (2008). Customer complaining: The role of tie strength and information control. *Journal of Retailing, 84*(2), 195-204.

Mittal, V., & Kamakura, W. A. (2001). Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research*, *38*(1), 131-142.

Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing*, *58*(3), 20-38.

Mostert, P., & De Meyer, C. (2010). Building customer relationships as retention strategy in the South African domestic passenger airline industry. *Acta Commercii, 10*, 27-42.

Mozer, M. C., Wolniewicz, R., Grimes, D. B., Johnson, E., & Kaushansky, H. (2000). Predicting subscriber dissatisfaction and improving retention in the wireless telecommunications industry. *Neural Networks, IEEE Transactions on, 11*(3), 690-696.

Narasimhan, C. (1984). A price discrimination theory of coupons. *Marketing Science, 3*(2), 128-147.

Nath, S. V., & Behara, R. S. (2003). *Customer Churn Analysis in the Wireless Industry: A Data Mining Approach.* Paper presented at the Proceedings-annual meeting of the Decision Sciences Institute.

Negi, R., & Ketema, E. (2010). Relationship marketing and customer loyalty: The Ethiopian mobile communications perspective. *International Journal of Mobile Marketing, 5*(1), 50-71.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, *43*(2), 204-211.

Nicholson, C. Y., Compeau, L. D., & Sethi, R. (2001). The role of interpersonal liking in building trust in long-term channel relationships. *Journal of the Academy of Marketing Science, 29*(1), 3-15.

Nitzan, I., & Libai, B. (2011). Social effects on customer retention. *Journal of Marketing, 75*(6), 24-38.

OECD Science, Technology and Industry Outlook 2014. Paris, France: OECD Publishing.

Owczarczuk, M. (2010). Churn models for prepaid customers in the cellular telecommunication industry using large data marts. *Expert Systems with Applications, 37*(6), 4710-4712.

Palmatier, R., Dant, R., Grewal, D., & Evans, K. (2006). Factors influencing the effectiveness of relationship marketing: A meta-analysis. *Journal of Marketing, 70*(4), 136-153.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL. *Journal of Retailing, 64*(1), 12-40.

Patterson, P. G. (2007). Demographic correlates of loyalty in a service context. *Journal of Services Marketing, 21*(2), 112-121.

Payne, A., & Frow, P. (2013). *Strategic Customer Management: Integrating Relationship Marketing and CRM.*: Cambridge University Press.

Payne, A., Storbacka, K., & Frow, P. (2008). Managing the co-creation of value. *Journal of the Academy of Marketing Science, 36*(1), 83-96.

Peng, L. Y., & Wang, Q. (2006). Impact of relationship marketing tactics (RMTs) on switchers and stayers in a competitive service industry. *Journal of Marketing Management, 22,* 25-59.

Petersen, J. A., McAlister, L., Reibstein, D. J., Winer, R. S., Kumar, V., & Atkinson, G. (2009). Choosing the right metrics to maximize profitability and shareholder value. *Journal of Retailing, 85*(1), 95-111.

Ping, R. A. (1993). The effects of satisfaction and structural constraints on retailer exiting, voice, loyalty, opportunism, and neglect. *Journal of Retailing, 69*(3), 320-352.

Prentice, C. (2014). Who stays, who walks, and why in high-intensity service contexts. *Journal of Business Research, 67*(4), 608-614.

Prins, R., & Verhoef, P. C. (2007). Marketing communication drivers of adoption timing of a new e-service among existing customers. *Journal of Marketing, 71*(2), 169-183.

Rappa, M. A. (2004). The utility business model and the future of computing services. *IBM Systems Journal, 43*(1), 32-42.

Reichheld, F. F. (1996). Learning from customer defections. *Harvard Business Review, 74*, 56-70.

Reichheld, F. F., & Sasser Jr, W. E. (1990). Zero defections: Quality comes to services. *Harvard Business Review, 68*(5), 105.

Reinartz, W., Thomas, J. S., & Kumar, V. (2005). Balancing acquisition and retention resources to maximize customer profitability. *Journal of Marketing*, *69*(1), 63-79.

Reinartz, W., & Ulaga, W. (2008). How to sell services more profitably. *Harvard Business Review, 86*(5), 90.

Reinartz, W. J., & Kumar, V. (2000). On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing. *Journal of Marketing*, 17-35.

Reinartz, W. J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing, 67*(1), 77-99.

Risselada, H., Verhoef, P. C., & Bijmolt, T. H. (2010). Staying power of churn prediction models. *Journal of Interactive Marketing, 24*(3), 198-208.

Roberts, K., Varki, S., & Brodie, R. (2003). Measuring the quality of relationships in consumer services: an empirical study. *European Journal of Marketing, 37*(1/2), 169-196.

Roschk, H., & Gelbrich, K. (2014). Identifying appropriate compensation types for service failures: A meta-analytic and experimental analysis. *Journal of Service Research, 17*(2), 195-211.

Roos, I., Edvardsson, B., & Gustafsson, A. (2004). Customer switching patterns in competitive and noncompetitive service industries. *Journal of Service Research, 6*(3), 256-271.

Ross, W. T., & Simonson, I. (1991). Evaluations of pairs of experiences: A preference for happy endings. *Journal of Behavioral Decision Making, 4*(4), 273-282.

Rust, R. T., Lemon, K. N., & Zeithaml, V. A. (2004). Return on marketing: Using customer equity to focus marketing strategy. *Journal of Marketing*, *68*(1), 109-127.

Rust, R. T., & Zahorik, A. J. (1993). Customer satisfaction, customer retention, and market share. *Journal of Retailing, 69*(2), 193-215.

Ryals, L. (2005). Making customer relationship management work: The measurement and profitable management of customer relationships. *Journal of Marketing*, *69*(4), 252-261.

Ryals, L. and Knox, S. (2001) Cross-functional issues in the implementation of relationship marketing through customer relationship management. *European Management Journal,19*(5), 534-542

Schumann, J. H., Wünderlich, N. V., & Evanschitzky, H. (2014). Spillover effects of service failures in coalition loyalty programs: The buffering effect of special treatment benefits. *Journal of Retailing, 90*(1), 111-118.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Schweidel, D. A., Fader, P. S., & Bradlow, E. T. (2008). Understanding service retention within and across cohorts using limited information. *Journal of Marketing, 72*(1), 82-94.

Seiders, K., & Berry, L. L. (1998). Service fairness: What it is and why it matters. *Academy of Management Executive, 12*(2), 8-20.

Shaffer, G., & Zhang, Z. J. (2002). Competitive one-to-one promotions. *Management Science, 48*(9), *48*(9), 1143-1160.

Sharma, P., Marshall, R., Alan Reday, P., & Na, W. (2010). Complainers versus non-complainers: A multi-national investigation of individual and situational influences on customer complaint behaviour. *Journal of Marketing Management, 26*(1-2), 163-180.

Sheth, J. N., & Parvatiyar, A. (1995). The evolution of relationship marketing. *International Business Review, 4*(4), 397-418.

Singer, J. D., & Willett, J. B. (2003). *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence.* New York, NY: Oxford University Press.

Smith, A. K., & Bolton, R. N. (1998). An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters: Paradox or Peril? *Journal of Service Research, 1*(1), 65-81.

Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research*, *36*(3), 356-372.

Snijders, T. A. (2011). *Multilevel analysis.* Berlin: Springer.

Spicer, J. (2005). *Making Sense of Multivariate Data Analysis: An Intuitive Approach.* Thousand Oaks, California : Sage Publications Inc.

Stauss, B., & Seidel, W. (2004). *Complaint Management: The Heart of CRM.* Ohio: Thompson/South-Western.

Tax, S. S., & Brown, S. W. (1998). Recovering and learning from service failure. *Sloan Management Review, 40*(1), 75-88.

Thompson, S. A., & Sinha, R. K. (2008). Brand communities and new product adoption: The influence and limits of oppositional loyalty. *Journal of Marketing, 72*(6), 65-80.

Toufaily, E., Ricard, L., & Perrien, J. (2013). Customer loyalty to a commercial website: Descriptive meta-analysis of the empirical literature and proposal of an integrative model. *Journal of Business Research, 66*(9), 1436-1447.

Vakratsas, D. (1998). Household cost effects on purchase timing decisions: Do demographics matter? *Journal of Consumer Marketing, 15*(1), 6-22.

Van Buuren, S., Brand, J. P., Groothuis-Oudshoorn, C., & Rubin, D. B. (2006). Fully conditional specification in multivariate imputation. *Journal of Statistical Computation and Simulation, 76*(12), 1049-1064.

Van den Bulte, C., & Iyengar, R. (2011). Tricked by truncation: Spurious duration dependence and social contagion in hazard models. *Marketing Science, 30*(2), 233-248.

Van den Poel, D., & Larivière, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research, 157*(1), 196-217.

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography, 16*(3), 439-454.

Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing, 68*(4), 106-125.

Verhoef, P. C. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, *67*(4), 30-45.

Wei, C.-P., & Chiu, I.-T. (2002). Turning telecommunications call details to churn prediction: a data mining approach. *Expert Systems with Applications, 23*(2), 103-112.

Weun, S., Beatty, S. E., & Jones, M. A. (2004). The impact of service failure severity on service recovery evaluations andpost-recovery relationships. *Journal of Services Marketing, 18*(2), 133-146.

Williams, T. D., Drake, M. F., & Moran, J. D. (1993). Complaint behaviour, price paid and the store patronized. *International Journal of Retail & Distribution Management, 21*(5), 3-9.

Winer, R. S. (2001). Customer relationship management: A framework for customer relationship management. *California Management Review. 43*(4), 89-105.

Wirtz, J., Xiao, P., Chiang, J., & Malhotra, N. (2014). Contrasting the drivers of switching intent and switching behavior in contractual service settings. *Journal of Retailing, 90*(4), 463-480.

Wong, K. K.-K. (2010). Fighting churn with rate plan right-sizing: A customer retention strategy for the wireless telecommunications industry. *The Service Industries Journal, 30*(13), 2261-2271.

Wong, K. K.-K. (2011). Using Cox regression to model customer time to churn in the wireless telecommunications industry. *Journal of Targeting, Measurement and Analysis for Marketing, 19*(1), 37-43.

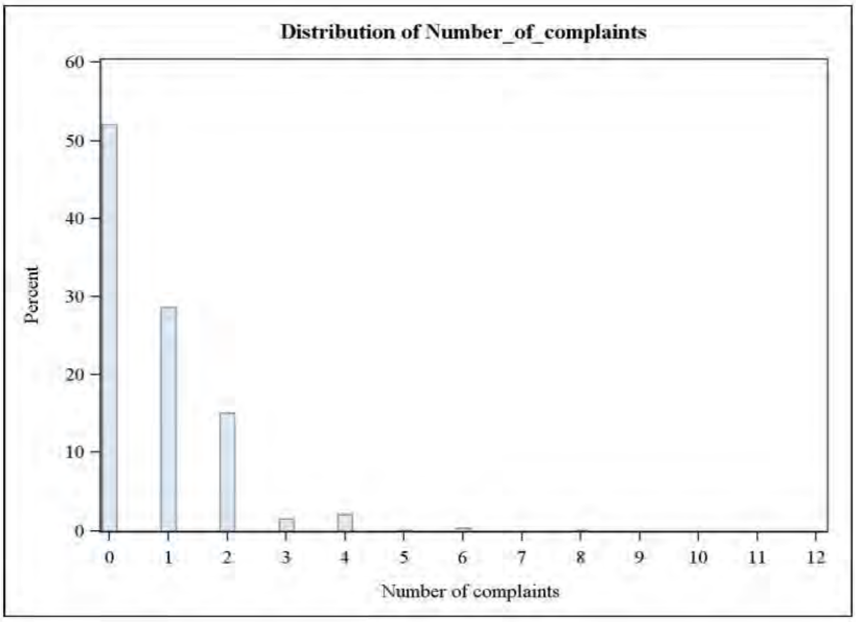
Wright, C., & Sparks, L. (1999). Loyalty saturation in retailing: exploring the end of retail loyalty cards? *International Journal of Retail & Distribution Management, 27*(10), 429-440.

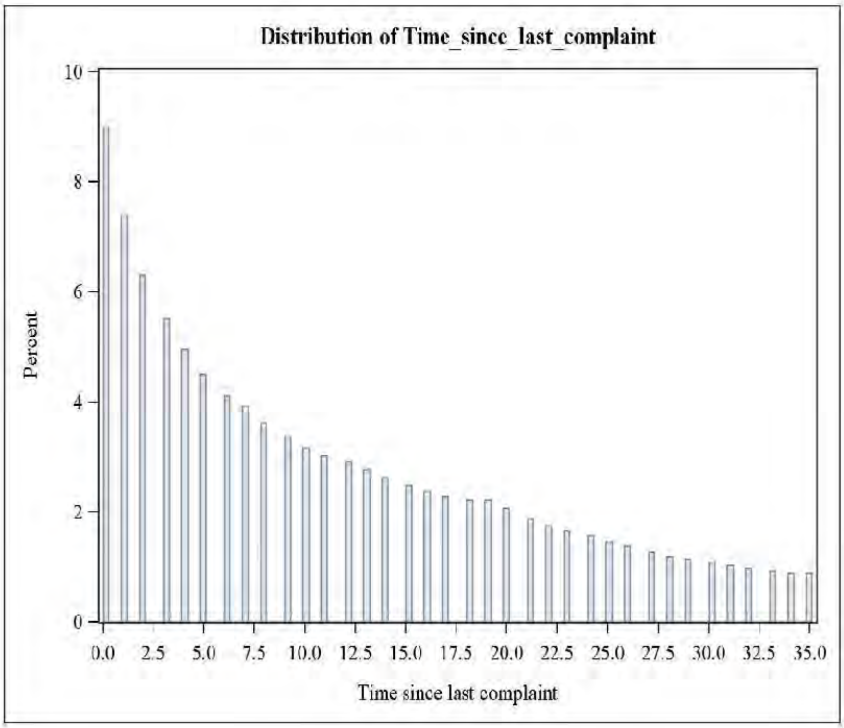
Yuan, M., & Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 68*(1), 49-67.

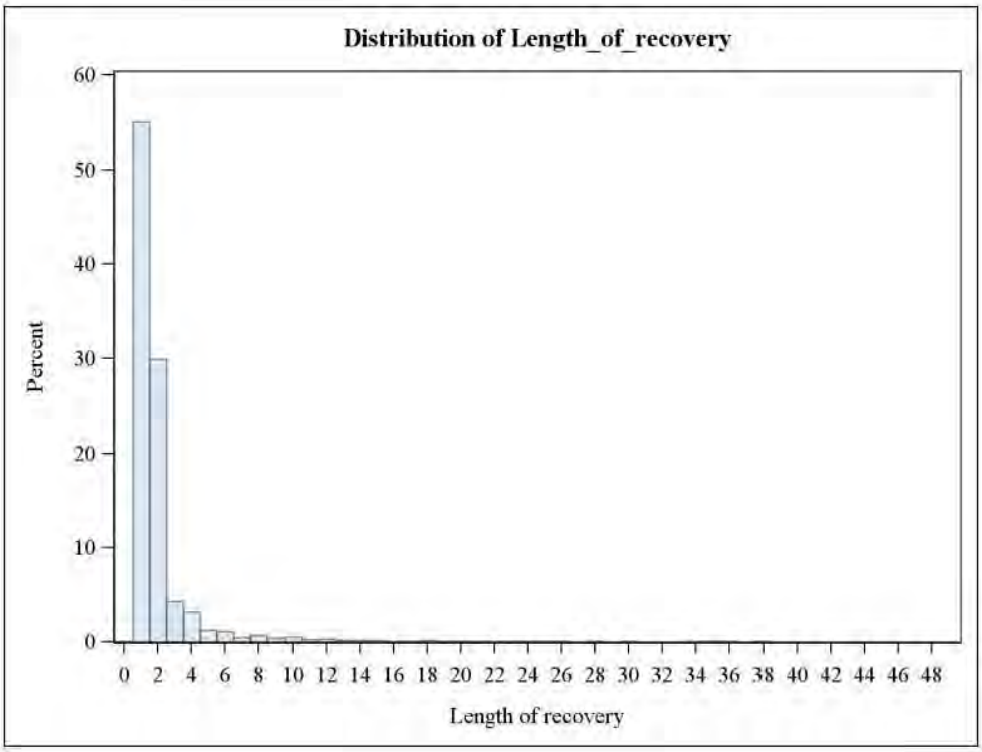
Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *The Journal of Marketing*, *60*(2), 31-46.

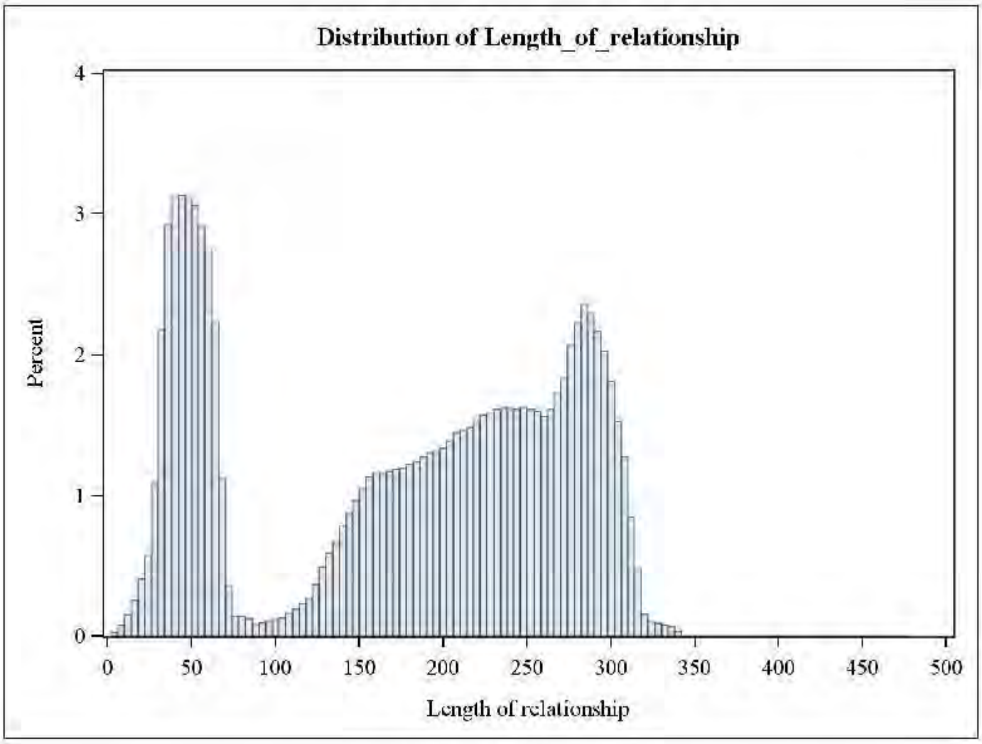
Appendix A: Graphical Examination of the Continuous Explanatory Variables for Voice and Internet

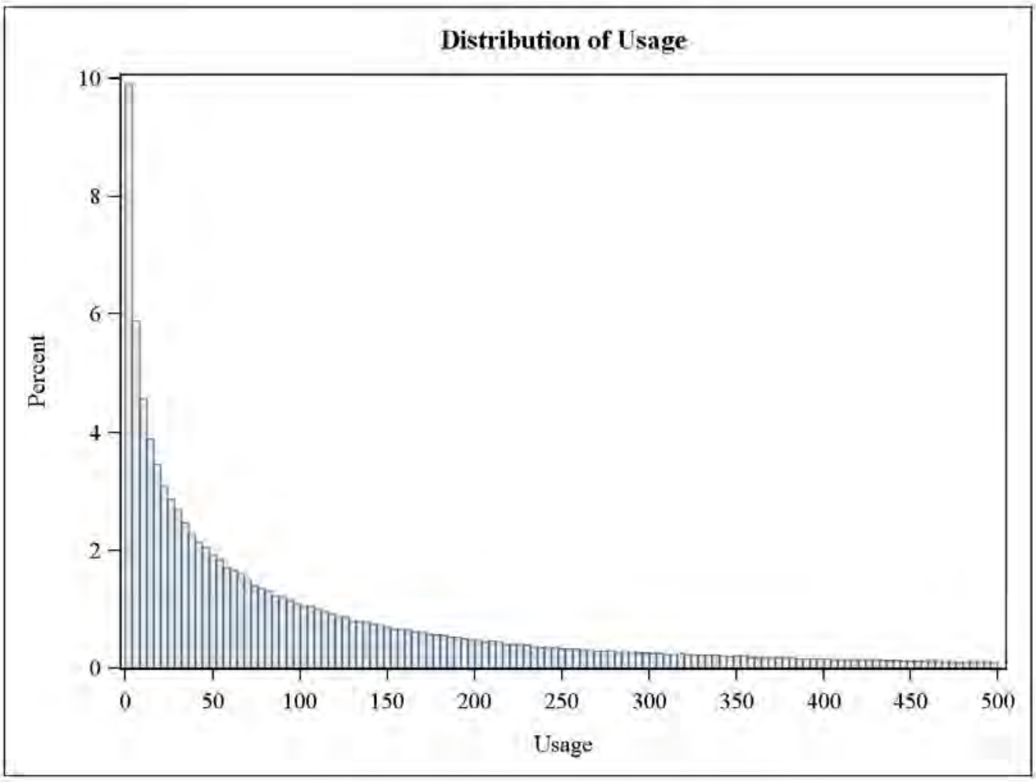
A1. Graphical Examination of the Continous Explanatory Variables for Voice Services

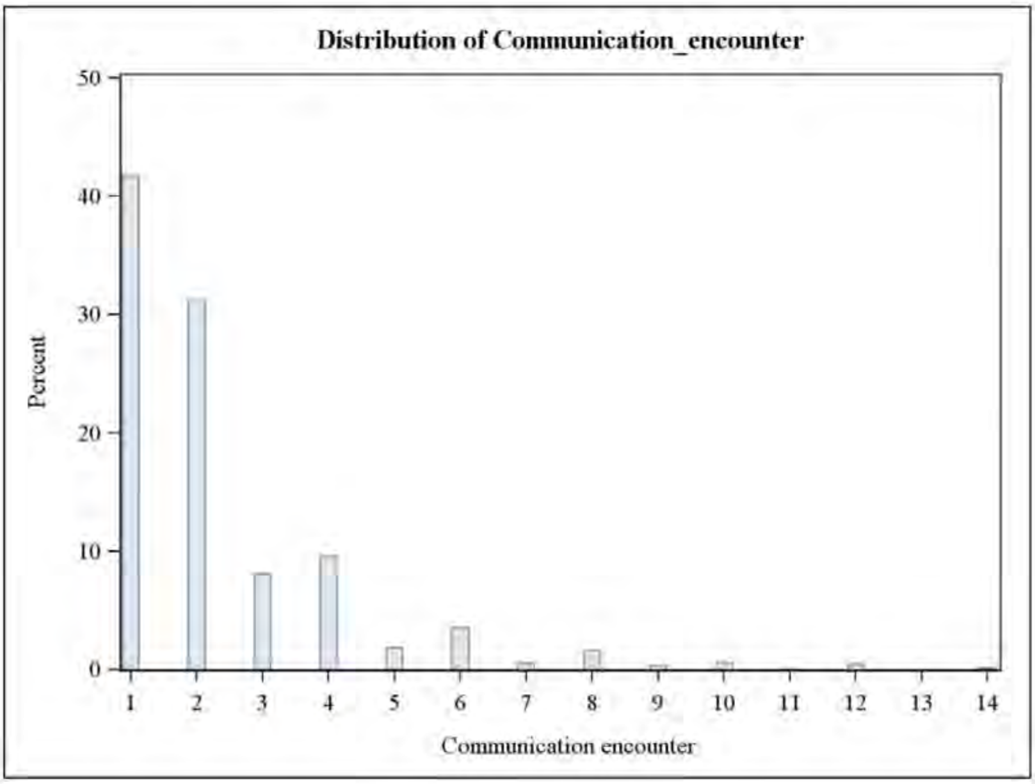


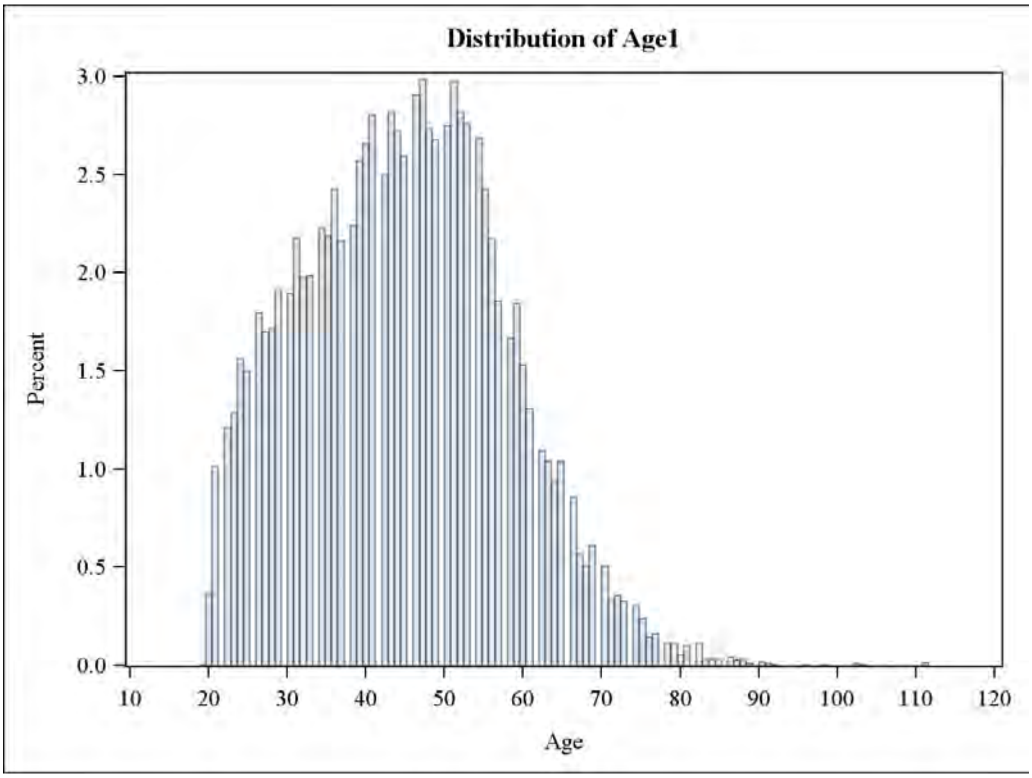



Distribution of Length_of_recovery
Percent
Length of recovery
0 10 20 30 40 50 60
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48


Distribution of Length_of_relationship
Percent
Length of relationship
0 1 2 3 4
0 50 100 150 200 250 300 350 400 450 500


Distribution of Usage
Percent
0
2
4
6
8
10
0
50
100
150
200
250
300
350
400
450
500
Usage

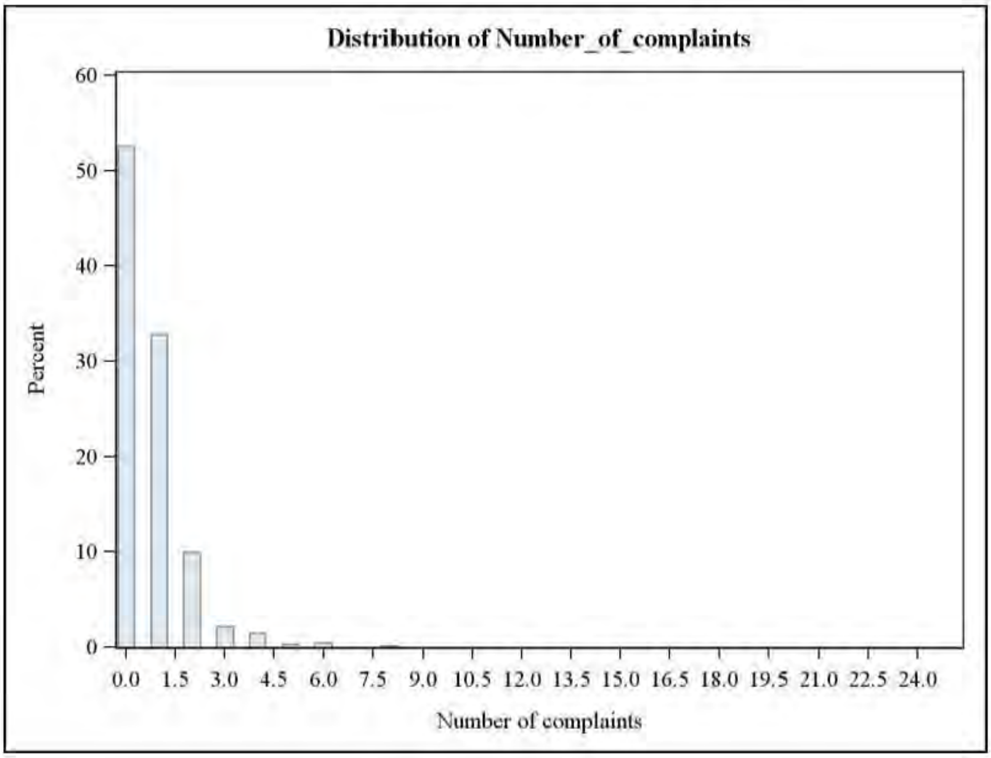

Distribution of Communication_encounter
Percent
0
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Communication encounter

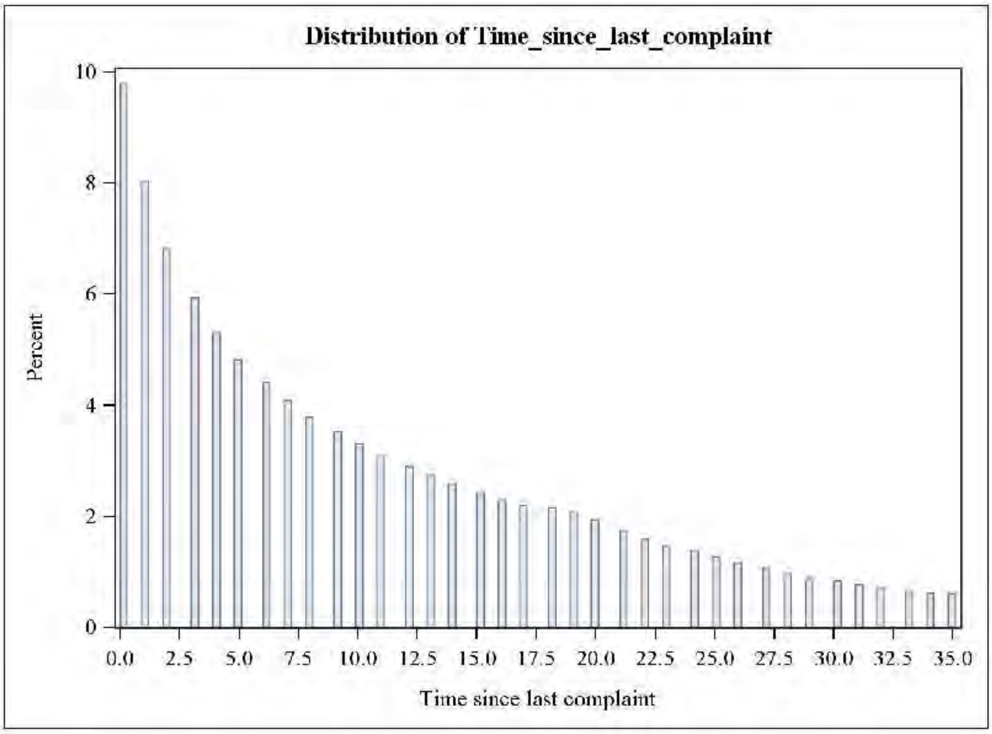


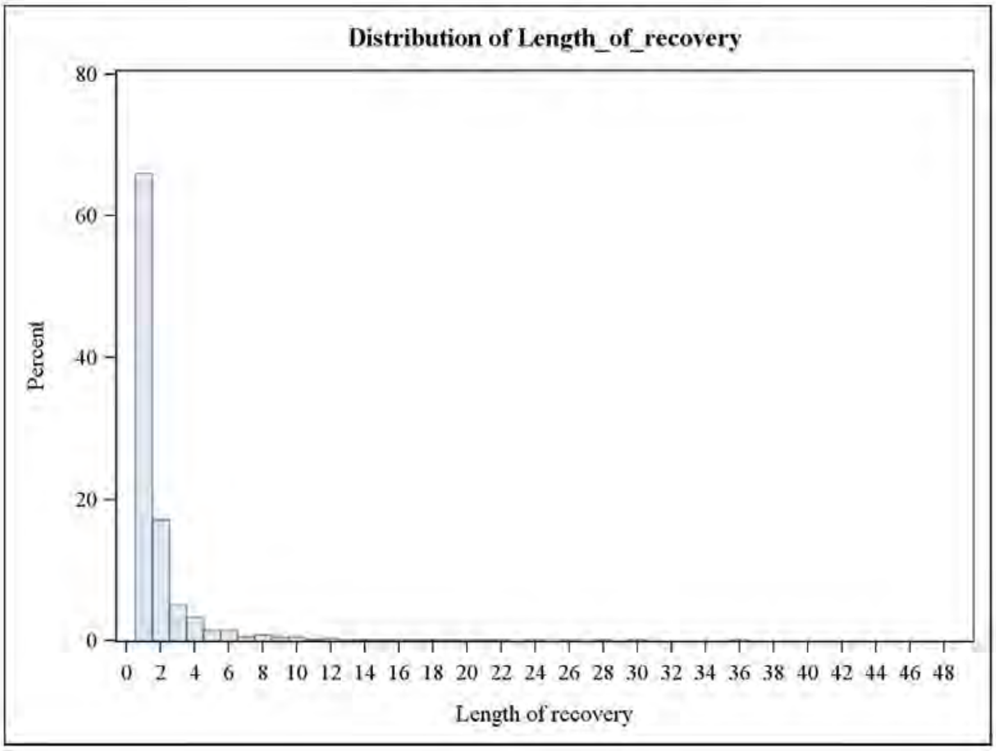
Note:

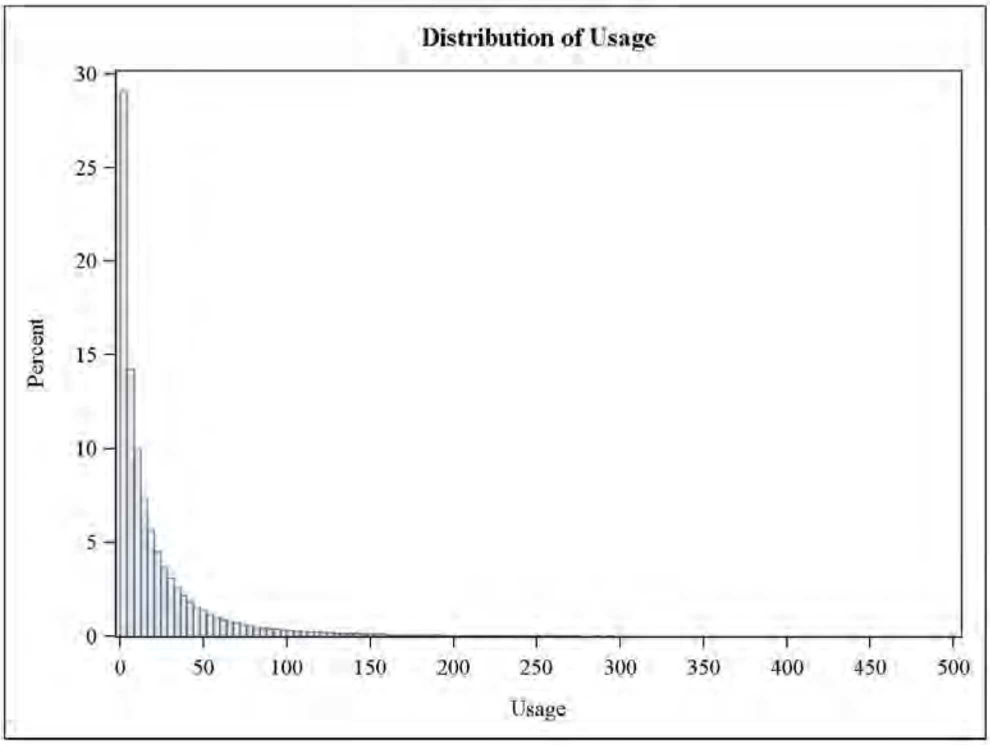
Age is operationalised as a categorical variable in the model estimation. As categorical variables it provides insight to customer defection behaviour by segments.

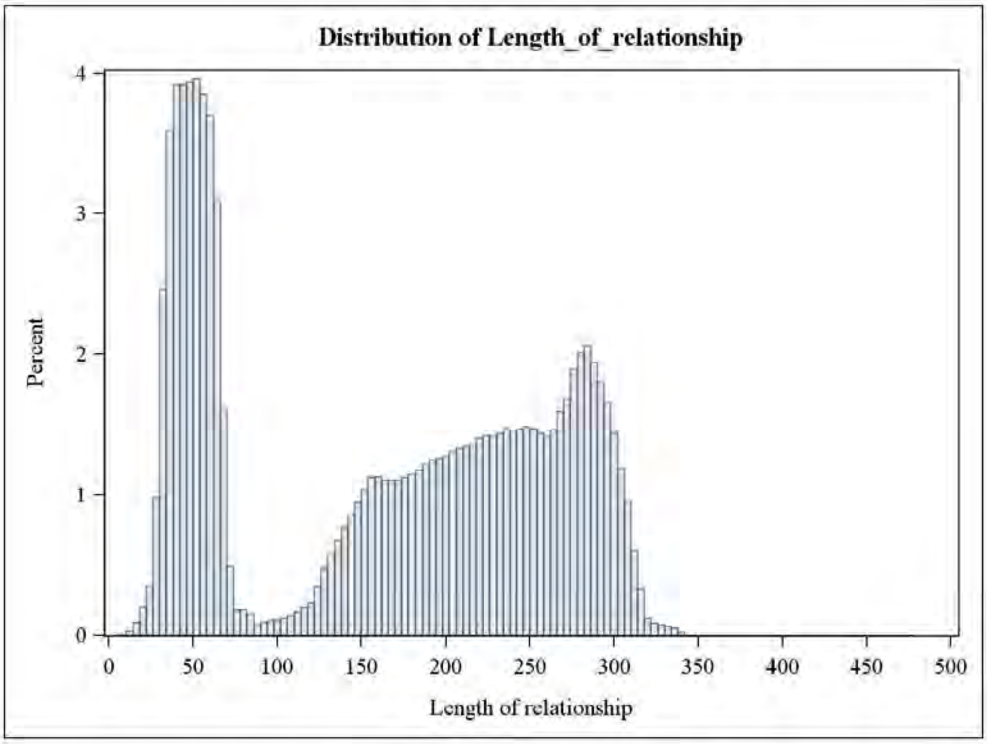
A2: Graphical Examination of the Explanatory Variables for Internet Services

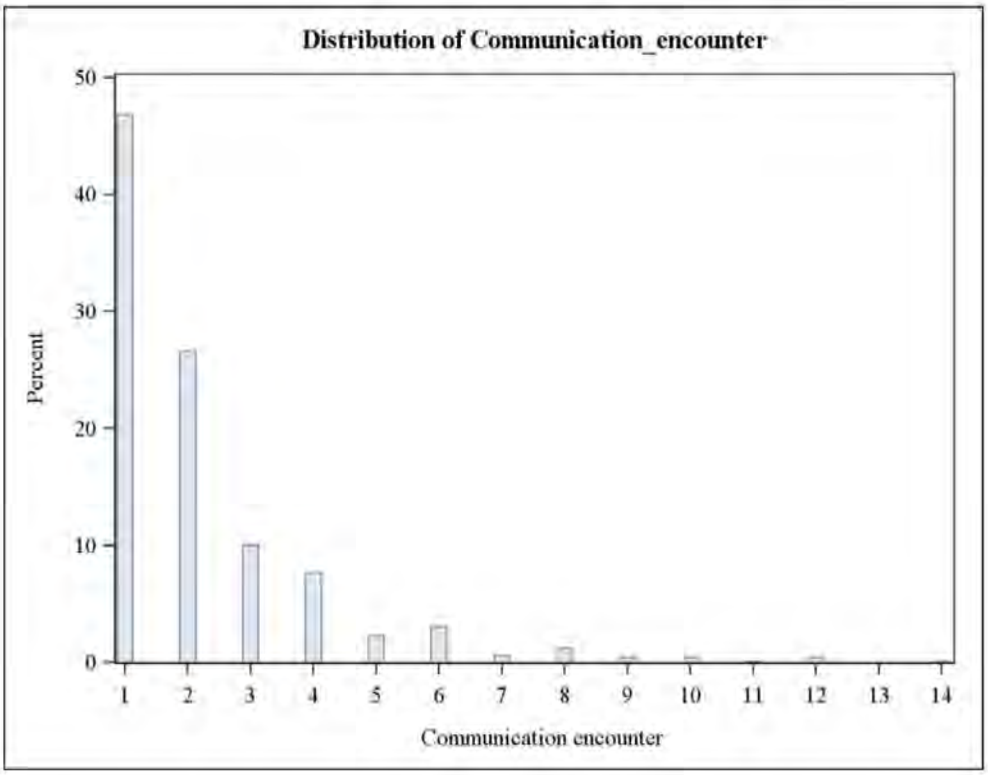


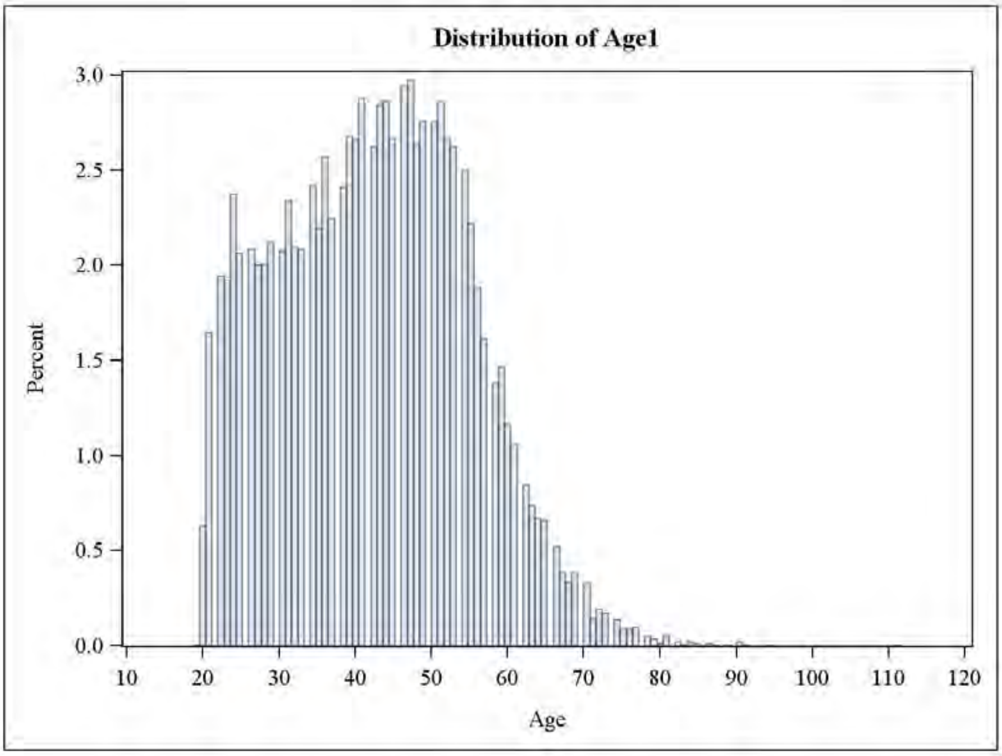



Distribution of Length_of_recovery
80
60
40
20
0
Percent
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44 46 48
Length of recovery


Distribution of Usage
30
25
20
15
10
5
0
Percent
0 50 100 150 200 250 300 350 400 450 500
Usage


Distribution of Length_of_relationship
Percent
0
1
2
3
4
0
50
100
150
200
250
300
350
400
450
500
Length of relationship


Distribution of Communication_encounter
Percent
0
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Communication encounter


Distribution of Age1
Percent
3.0
2.5
2.0
1.5
1.0
0.5
0.0
10
20
30
40
50
60
70
80
90
100
110
120
Age

Appendix B: Correlation Matrix for Voice and Internet

B1: Correlation Analysis for Voice Services

Pearson Correlation Coefficients/ Prob > |R| under H_0: Rho=0 /Number of Observations

Explanatory variables	Number of complaints	Usage	Time since last complaint	Length of relationship	Communication encounter	Length of recovery
Number of complaints	1.000	0.018	-0.176	0.012	0.640	0.229
Usage	0.018	1.000	0.113	0.110	0.026	0.004
Time since last complaint	-0.176	-0.113	1.000	0.080	-0.226	-0.090
Length of relationship	0.012	0.110	0.080	1.000	0.011	0.00094*
Communication encounter	0.640	0.026	-0.226	0.011	1.000	0.288
Length of recovery	0.229	0.004	-0.090	0.00094*	0.288	1.000

Note:

All variables are significant at p<.0001

* p=.4594

B2: Correlation Analysis for Internet Services

Pearson Correlation Coefficients/ Prob > |R| under H_0: Rho=0 /Number of Observations

Explanatory variables	Number of complaints	Usage	Time since last complaint	Length of relationship	Communication encounter	Length of recovery
Number of complaints	1.000	0.058	-0.162	0.031	0.639	0.249
Usage	0.058	1.000	-0.002	0.032	0.075	0.029
Time since last complaint	-0.162	-0.002	1.000	-0.00208*	0.214	-0.086
Length of relationship	0.031	0.032	-0.00208*	1.000	0.035	0.010
Communication encounter	0.639	0.075	-0.214	0.035	1.000	0.288
Length of recovery	0.249	0.029	-0.086	0.010	0.288	1.000

Note:

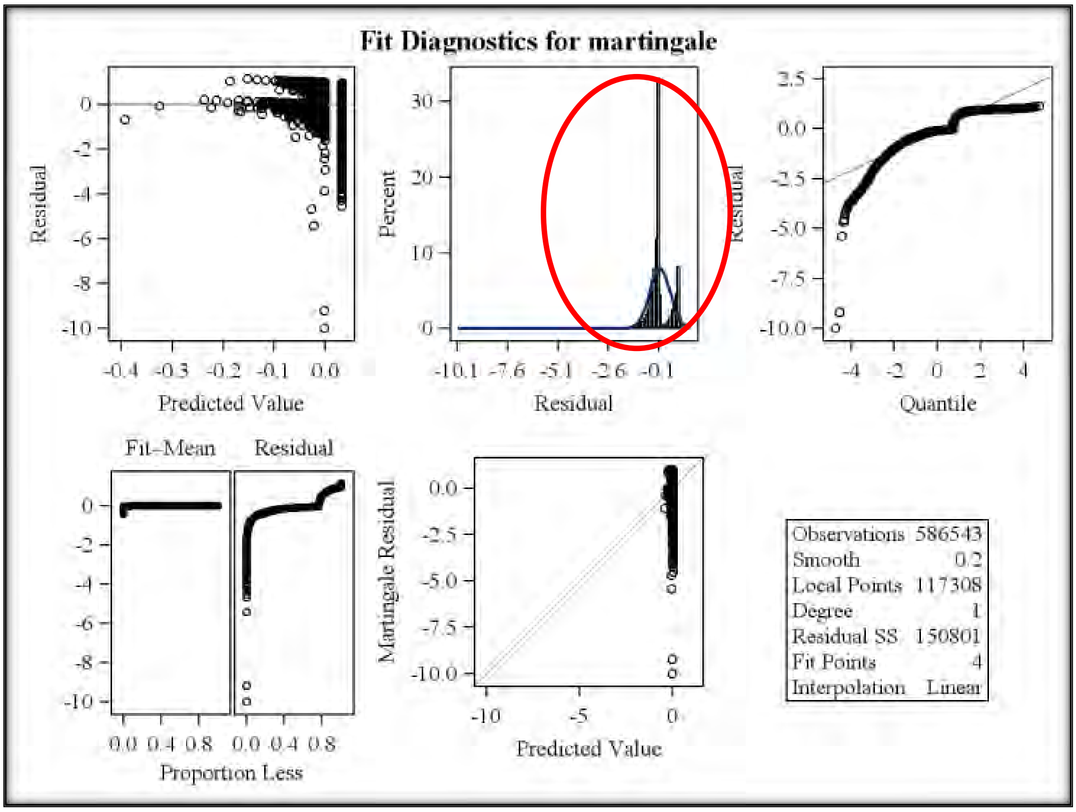
All variables are significant at p<.0001

* p=0.1106

Appendix C: Proportionality Assumption Results for Cox PH Model

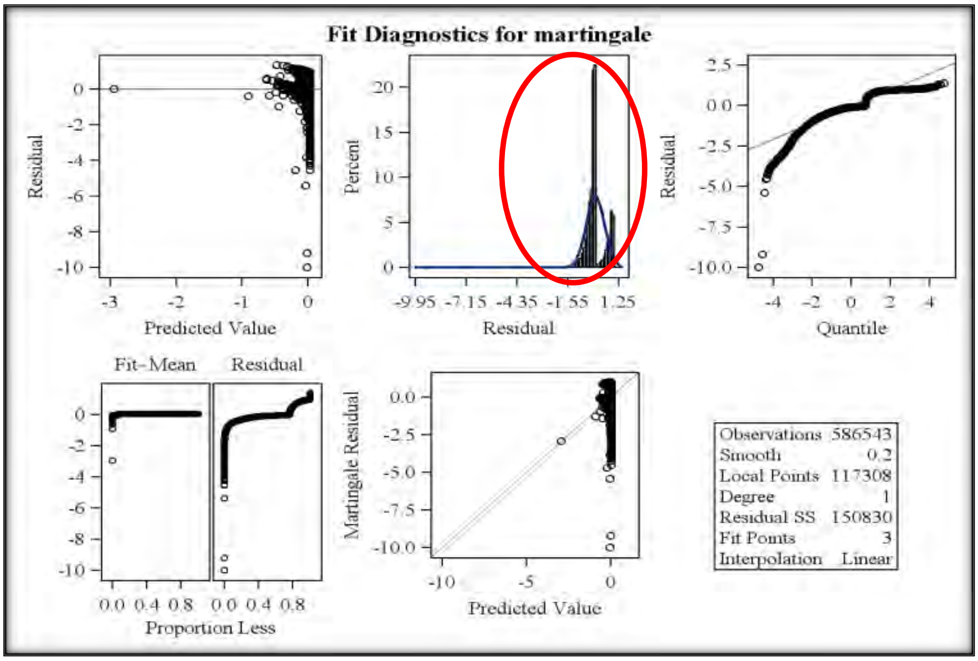
C1: Schoenfeld Residuals Test for Voice Services

Variable: Communication Encounter[41]

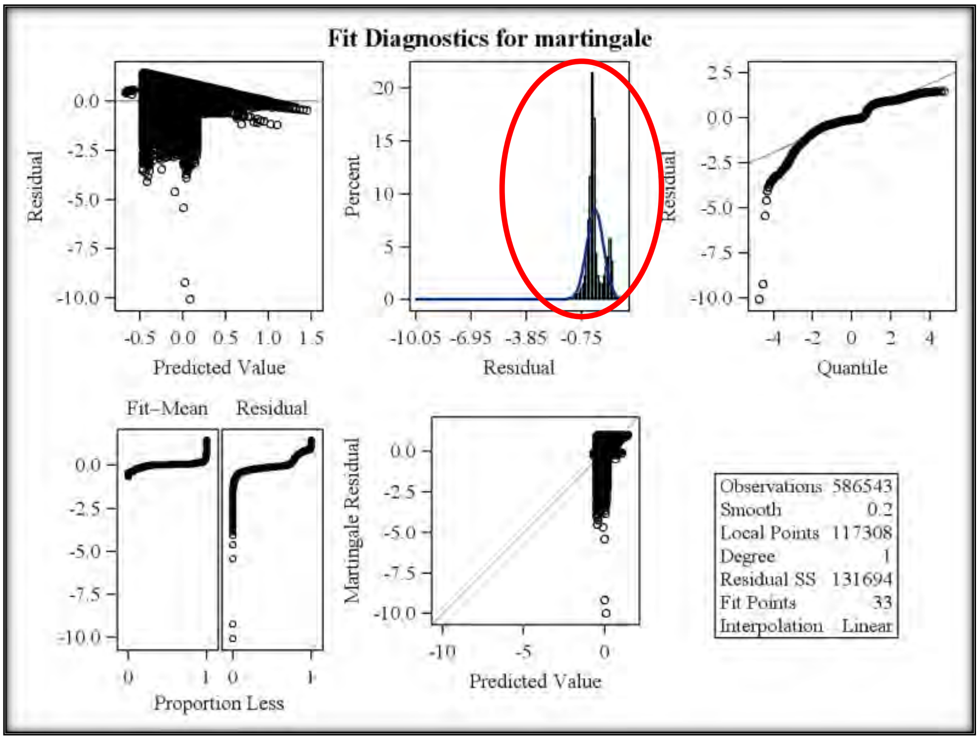


[41] The martingale residuals graphs are used to determine the correct form through which an Independent Variable (IV) affects the Dependent Variable (DV). That is whether the form should be linear (y=f(x)) or square (y=f(x^2) or logarithmic(y=f(ln(x)). The third and fourth graphs in the set of graphs are the most important one since they display the presence or absence of a functional form. In our case the residuals are displaying a random normal pattern and hence we can reject the claim that the form in which they have entered the equation is incorrect. For completeness, all the diagrammatic outputs of the statistics are included; however the important one to consider is the residual plot because it shows if the effect of the independent variable on the dependent variable is seen to vary with time. This is highlighted with a circle in these following images of the tests/statistics. This comment applies to all the following graphs presented in Appendix C.

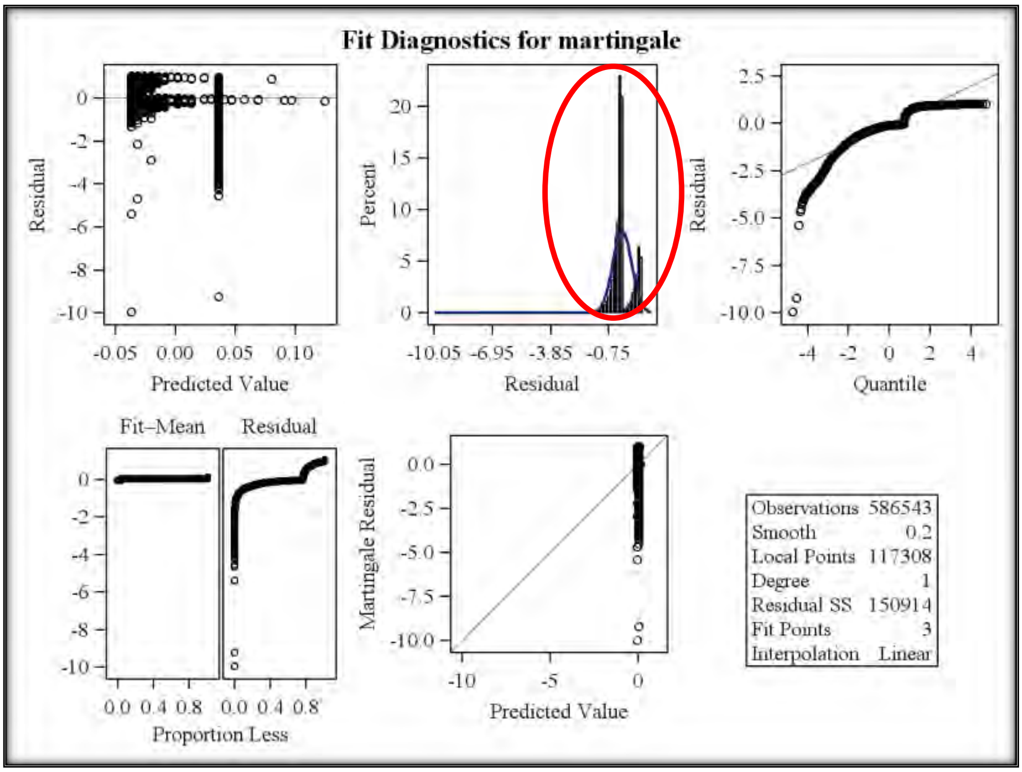
Variable: Length of Recovery



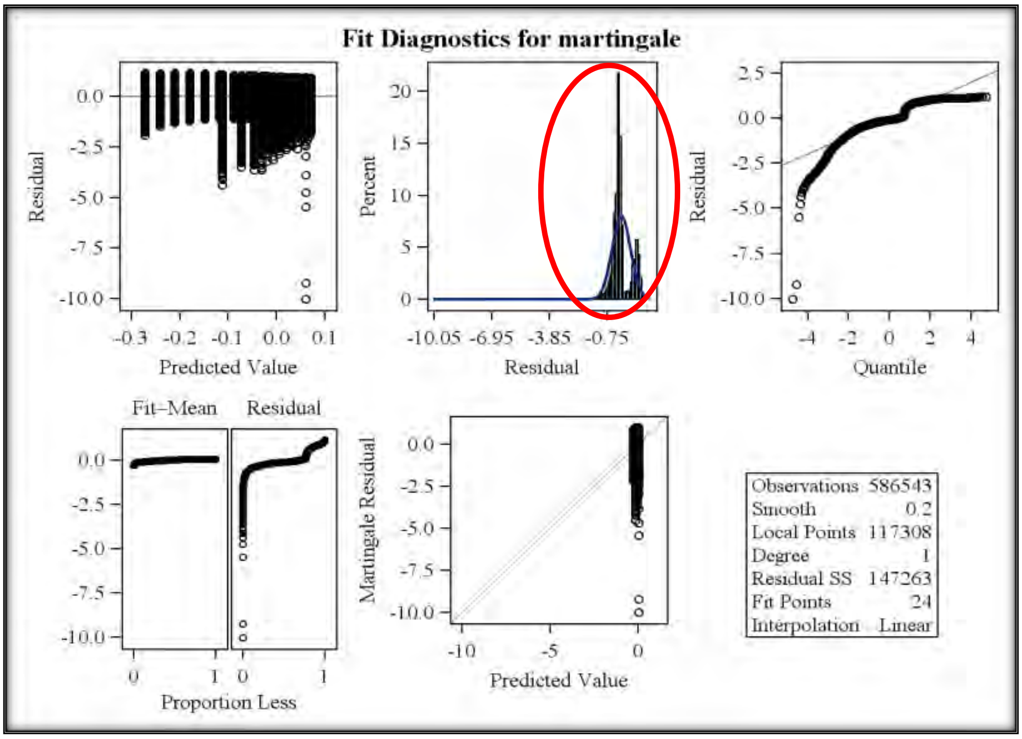
Variable: Length of Relationship



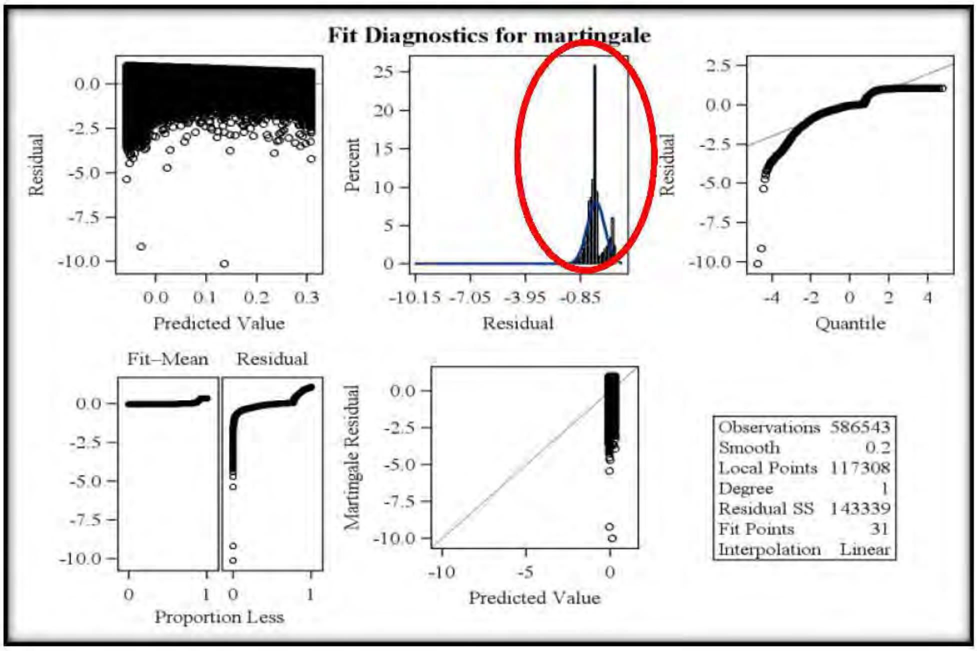
Variable: Number of Complaints



Variable: Time Since Last Complaint

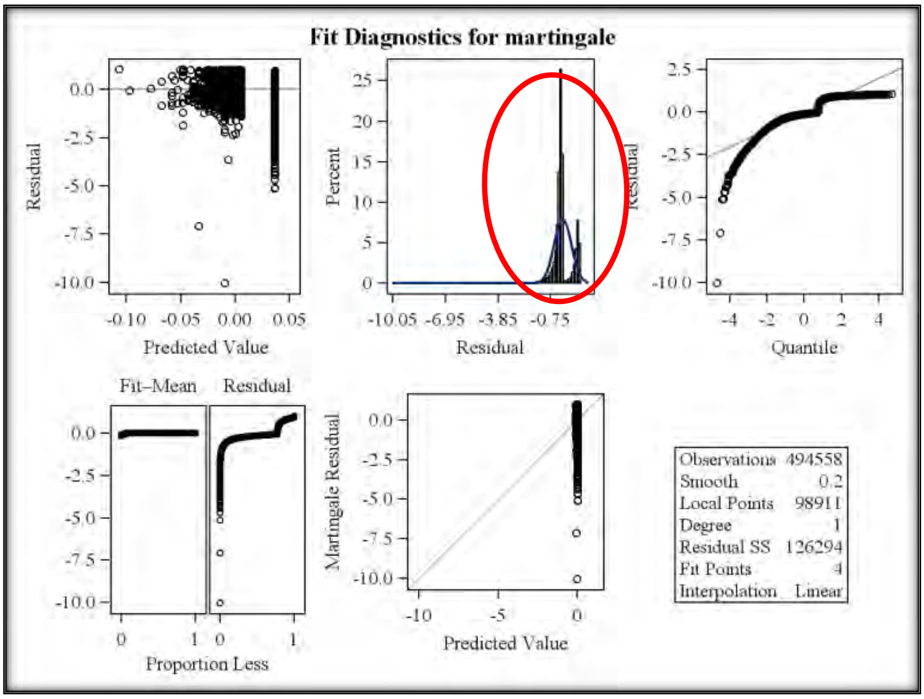


Variable: Usage

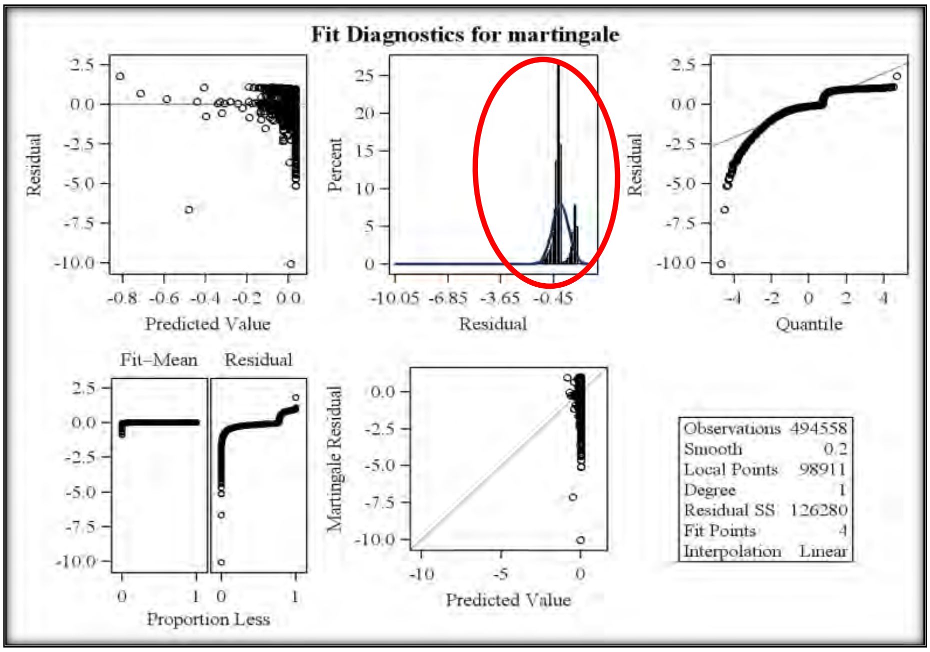


C2: Schoenfeld Residuals Test for Internet Services

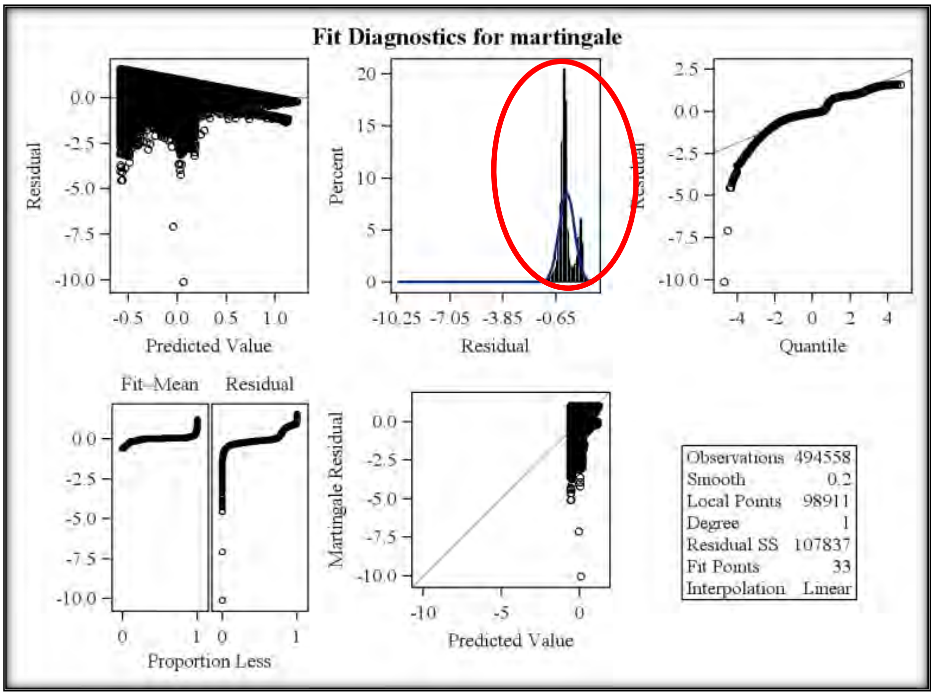
Variable: Communication Encounter



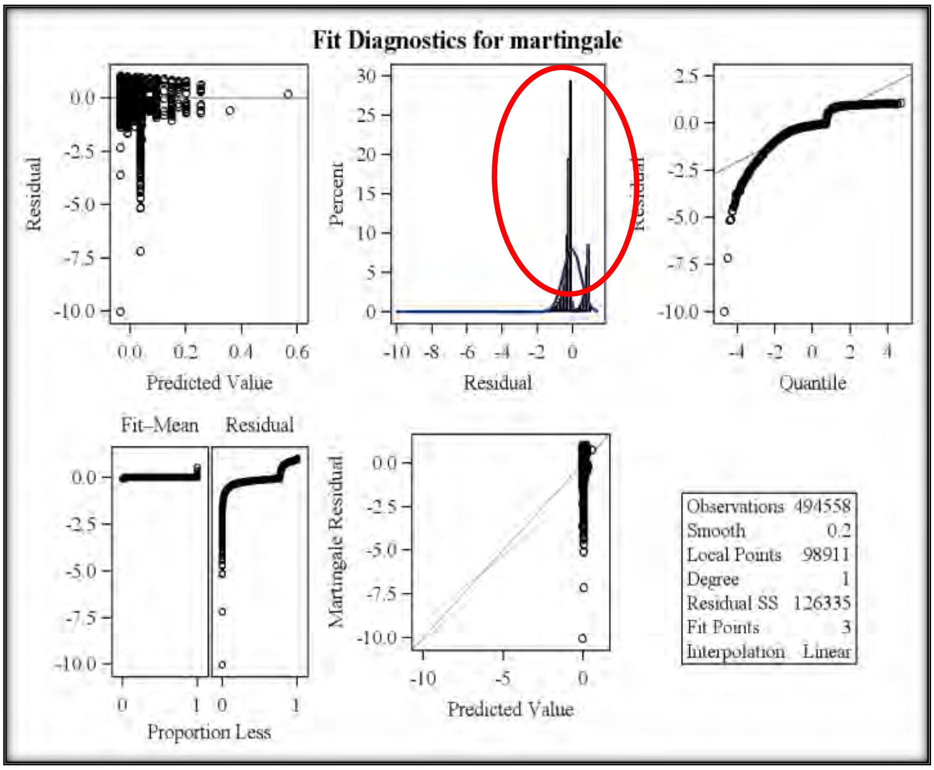
Variable: Length of Recovery



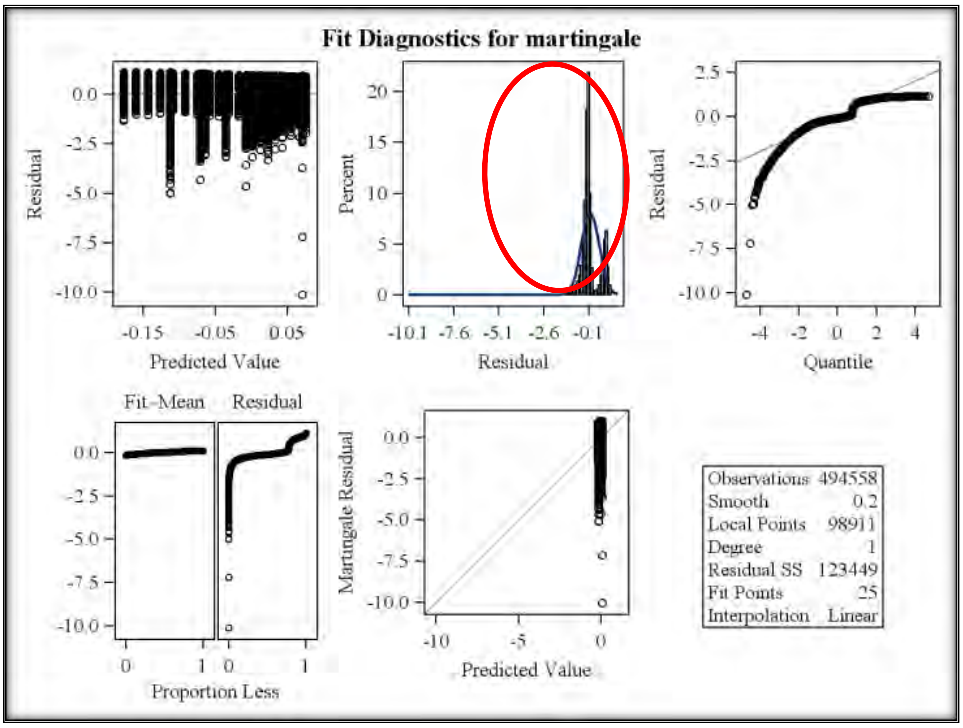
Variable: Length of Relationship



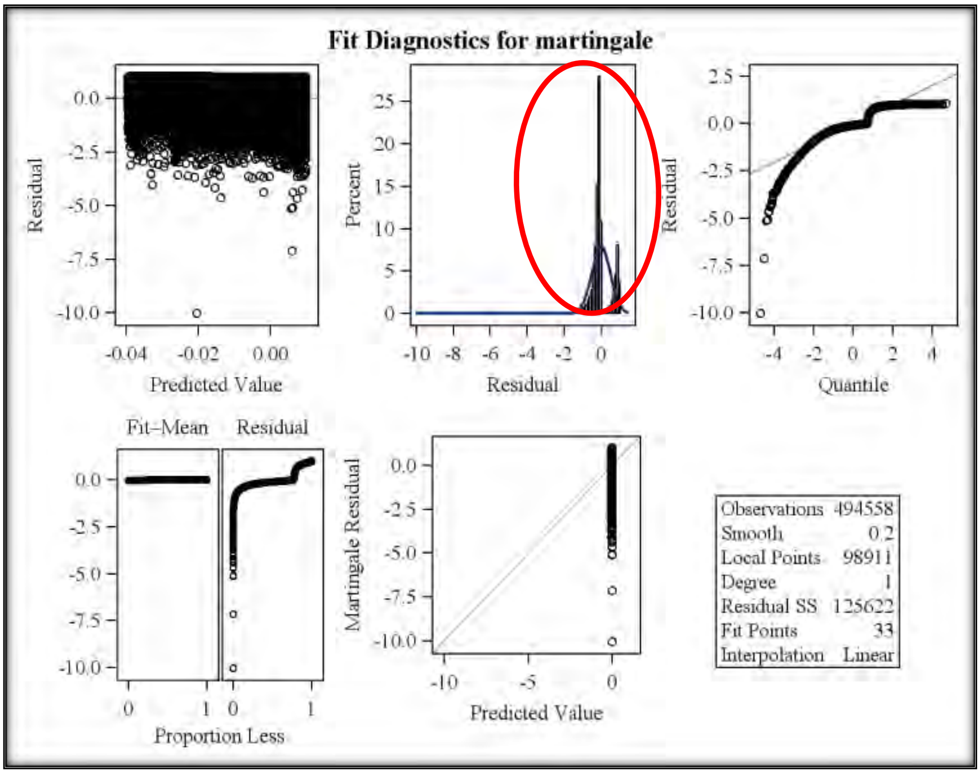
Variable: Number of Complaints



Variable: Time since Last Complaint



Variable: Usage



Appendix D: Summary Results for Three Models Voice and Internet

D1. Model Results for Voice Services

Variable	Model 1 (Baseline)		Model 2 (Comprehensive)		Model 3 (Interaction)	
	Parameter Estimates	Hazard Ratio	Parameter Estimates	Hazard Ratio	Parameter Estimates	Hazard Ratio
Age group 1 (19-25) (reference group)						
Group 2 (26-40)	0.485	1.623	0.188	1.206	0.027 ***	1.027
Group 3 (41-55)	0.498	1.645	0.376	1.456	0.285	1.33
Group 4 (over 55)	0.327	1.386	0.281	1.324	0.340	1.405
Gender male (reference group)	0.026	1.026	-0.006 ***	0.994	-0.006 ***	0.994
Household income low (reference group)						
Household income (medium)	-0.674	0.510	0.277	1.32	0.241	1.273
Household income (high)	-0.419	0.658	0.740	2.097	0.175 *	1.191
Household size 01-02 (reference group)						
Household size 03-04	0.034	1.035	-0.075	0.928	-0.076	0.927
Household size 05-06	-0.078	0.925	-0.096	0.908	-0.097	0.908
Household size over 6	-0.082	0.921	-0.083	0.92	-0.085	0.919
Usage	0.000	1.000	0.002	1.002	0.002	1.002
Length of relationship			-0.004	0.996	-0.004	0.996
Communication encounter			0.041	1.041	0.042	1.043
Complainants (reference group)			1.081	3.270	1.006	2.735
Number of complaints			-0.045	0.956	-0.050	0.951
Time since last complaints			-0.040	0.961	-0.040	0.961
Severity none (reference group)						
High			-0.847	0.429	-0.858	0.424
Low			1.063	2.894	1.095	2.989
Length of recovery			0.005	1.005	0.006	0.994
IncomeGroup (medium) * age (26-40)					0.179	1.196
IncomeGroup (medium) * age (41-55)					0.060 ***	1.062
IncomeGroup (medium) * age (>55)					-0.065 **	0.937
IncomeGroup (high) * age (26-40)					0.712	2.038
IncomeGroup (high) * age (41-55)					0.589	1.802
IncomeGroup (high) * age (>55)					0.493	1.637
Length of recovery * severity (high)					0.014	1.014
Usage * age (26-40)					0.000 *	1.000
Usage * age (41-55)					0.000 ***	1.000
Usage * age (>55)					0.000 ***	1.000
-2 LOG L		3,122,388		2,812,230		2,812,004
AIC		3,122,038		2,812,264		2,812,060
SBC		3,122,167		2,812,428		2,812,330

Note:

1. All explanatory variables are significant at p <.0001 unless otherwise stated in *
2. All marked with * is significant at p=0.05, ** is significant at p=0.01; *** is not significant.

D2. Model Results for Internet Services

Variable	Model 1 (Baseline)		Model 2 (Comprehensive)		Model 3 (Interaction)	
	Parameter Estimates	Hazard Ratio	Parameter Estimates	Hazard Ratio	Parameter Estimates	Hazard Ratio
Age group 1 (19-25) (reference group)						
Group 2 (26-40)	-0.491	0.612	0.181	1.198	0.384	1.468
Group 3 (41-55)	-0.271	0.763	0.389	1.475	0.545	1.725
Group 4 (over 55)	-0.067	0.934	0.463	1.589	0.529	1.697
Gender male (reference group)	-0.101	0.896	-0.120	0.887	-0.121	0.886
Household income low (reference group)						
Household income (medium)	-0.918	0.427	0.227	1.255	0.380	1.462
Household income (high)	-0.628	0.534	0.859	2.36	0.484	1.623
Household size 01-02 (reference group)						
Household size 03-04	0.011	1.011	-0.064	0.938	-0.062	0.94
Household size 05-06	-0.066	0.936	-0.133	0.876	-0.132	0.876
Household size over 6	-0.122	0.886	-0.061	0.941	-0.060	0.942
Usage			-0.004	0.996	-0.004	0.996
Length of relationship			-0.002	0.998	-0.002	0.998
Communication encounter			0.026	1.027	0.022	1.022
Complainants (reference group)			0.421	1.523	1.351	3.862
Number of complaints			-0.160	0.852	-0.171	0.843
Time since last complaints			-0.040	0.961	-0.024	0.976
Severity none (reference group)						
High			-1.028	0.358	-1.048	0.351
Low			1.542	4.672	1.557	4.745
Length of recovery			0.005	1.005	0.004 *	1.004
IncomeGroup (medium) * age (26-40)					-0.258	0.773
IncomeGroup (medium) * age (41-55)					-0.227	0.797
IncomeGroup (medium) * age (>55)					-0.110 *	0.896
IncomeGroup (high) * age (26-40)					0.245	1.278
IncomeGroup (high) * age (41-55)					0.342	1.408
IncomeGroup (high) * age (>55)					0.416	1.516
Length of recovery * severity (high)					0.011	1.011
Usage * age (26-40)					-0.001 *	0.999
Usage * age (41-55)					-0.001 **	0.999
Usage * age (>55)					-0.001 *	0.999
-2 LOG L		3,492,012		3,473,028		3,472,732
AIC		3,492,038		3,473,062		3,472,788
SBC		3,492,167		3,473,230		3,473,064

Note:

1. *All explanatory variables are significant at p <.0001 unless otherwise stated in
2. All marked with * is significant at p=0.05, ** is significant at p=0.01; *** is not significant

Appendix E: Standard Error Information for Voice and Internet Services

E1: Standard Error Information for Voice Services

Explanatory variable	Model 1 (Baseline)	Model 2 (Comprehensive)	Model 3 (Interaction)
Age group 1 (19-25) (reference group)			
Group 2 (26-40)	0.014	0.01426	0.0369
Group 3 (41-55)	0.010	0.01533	0.0356
Group 4 (over 55)	0.008	0.01708	0.0348
Gender male (reference group)	0.00588	0.00580	0.0062
Household income low (reference group)			
Household income (medium)	0.0132	0.0125	0.0294
Household income (high)	0.0068	0.0137	0.0813
Household size 01-02 (reference group)			
Household size 03-04	0.0119	0.0075	0.0077
Household size 05-06	0.0111	0.0085	0.0085
Household size over 6	0.0114	0.0126	0.0126
Usage	0.000	0.000025	0.0001
Length of relationship		0.000041	0.0000
Communication encounter		0.00294	0.0048
Complainants (reference group)		0.0091	0.0154
Number of complaints		0.0106	0.0121
Time since last complaints		0.000425	0.0003
Severity none (reference group)			
High		0.01430	0.0163
Low		0.02701	0.0313
Length of recovery		0.00035	0.0036
IncomeGroup (medium) * age (26-40)			0.0396
IncomeGroup (medium) * age (41-55)			0.0379
IncomeGroup (medium) * age (>55)			0.0370
IncomeGroup (high) * age (26-40)			0.0860
IncomeGroup (high) * age (41-55)			0.0850
IncomeGroup (high) * age (>55)			0.0856
Length of recovery * severity (high)			0.0037
Usage * age (26-40)			0.0001
Usage * age (41-55)			0.0001
Usage * age (>55)			0.0001

Source: PHREG output for Voice Services

E2: Standard Error Information for Internet Services

Explanatory variable	Model 1 (Baseline)	Model 2 (Comprehensive)	Model 3 (Interaction)
Age group 1 (19-25) (reference group)			
Group 2 (26-40)	0.014	0.01103	0.0283
Group 3 (41-55)	0.010	0.01223	0.0295
Group 4 (over 55)	0.008	0.01402	0.0308
Gender male (reference group)	0.00594	0.00569	0.0058
Household income low (reference group)			
Household income (medium)	0.0122	0.0112	0.021
Household income (high)	0.0062	0.0473	0.051
Household size 01-02 (reference group)			
Household size 03-04	0.0115	0.0069	0.007
Household size 05-06	0.0108	0.0077	0.008
Household size over 6	0.0111	0.0116	0.012
Usage	0.000069	0.0000765	0.0002361
Length of relationship		0.000040	0.0000396
Communication encounter		0.00477	0.00502
Complainants (reference group)		0.0050	0.0085
Number of complaints		0.0128	0.01493
Time since last complaints		0.000278	0.000376
Severity none (reference group)			
High		0.01870	0.01900
Low		0.02208	0.02225
Length of recovery		0.00041	0.00102
IncomeGroup (medium) * age (26-40)			0.02996
IncomeGroup (medium) * age (41-55)			0.0308
IncomeGroup (medium) * age (>55)			0.03214
IncomeGroup (high) * age (26-40)			0.05604
IncomeGroup (high) * age (41-55)			0.05618
IncomeGroup (high) * age (>55)			0.05781
Length of recovery * severity (high)			0.00152
Usage * age (26-40)			0.0002703
Usage * age (41-55)			0.0002615
Usage * age (>55)			0.0003269

Source: PHREG output for Voice Services

Appendix F: Model Evaluation

In general there are two approaches to evaluate the results, first by either Akaike's Information Criterion (AIC) or Schwarz-Bayes Criterion (SBC) (Akaike, 1981; Schwarz, 1978) and models that produce smaller value indicates strong evidence in terms of information loss. In addition, the likelihood ratio test (LRT) also compares the goodness of fit of two models, one of which (the null model) is a special case of the alternative model (Berger, Schäfer, & Ulm, 2003).

Based on the models estimate presented in Tables 5.1 and 5.2 the smallest AIC or the SBC is found for the comprehensive models with interactions. Further, LRT test detail provides support for the comprehensive models for both voice and Internet services. Table F1 and Table F2 provides the summary of AIC and SBC values for all three models and the LRT test for voice and Internet services[42].

Table F1 shows the summary of model evaluation for voice services. From the information above, it can be inferred that a smaller AIC and SBC is from the interactions model. For voice product, the LRT comparing the baseline and comprehensive models has chi-squared statistic of 42,540 with degree of freedom 7, so $p < 0.0001$ in favour of the comprehensive model. Subsequently, LRT comparing the comprehensive and

[42] The information is gathered from the final table of output which includes the chi-square values and p-values for the pairwise comparisons.

comprehensive with interactions models has chi-squared statistic of 740 with degree of freedom 10, so p < 0.0001 in favour of the comprehensive model with interactions.

Table F1: Summary of Model Evaluation for Voice Services

Model	AIC	SBC	-2 LOG L	DF	Chi square statistic
Model 1- Baseline	3,122,410	3,122,517	3,122,388		-
Model 2-Comprehensive	2,812,264	2,812,428	2,812,230	7	42,540
Model 3- Interaction	2,812,060	2,812,330	2,812,003	10	740

Table F2 shows the summary of model evaluation for Internet services. The AIC and SBC is smaller for the interaction models. Conversely, for Internet product the LRT comparing the baseline and comprehensive models has chi-squared statistic of 39,700 with degree of freedom 7, so p < 0.0001 in favour of the comprehensive model. Additionally, the LRT comparing the comprehensive and comprehensive with interactions models has chi-squared statistic of 17,975 with degree of freedom 10, so p < 0.0001 in favour of the comprehensive model with interactions.

Table F2: Summary of Model Evaluation for Internet Services

Model	AIC	SBC	-2 LOG L	DF	Chi square statistic
Model 1- Baseline	3,492,038	3,492,167	3,492,012		-
Model 2-Comprehensive	3,473,062	3,473,230	3,473,028	7	39,700
Model 3- Interaction	3,472,788	3,473,064	3,472,732	10	17,975

In both the voice and Internet services, there is evidence to favour the interactions models over the base line and comprehensive models. At the same time, LRT test detail provides support for the comprehensive models for both voice and Internet services. The final model for both Internet and voice services is derived mostly from the comprehensive with interactions models and the results are interpreted from Model three, while Model one and Model two are used for validation of our interpretation.