Understanding the influence basic service and value-added service usage has on customer churn in Pay-TV subscription services

Marc-André Gosselin

A Thesis
in
The Department
of
Marketing

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science Administration (Marketing) at
Concordia University
Montreal, Quebec, Canada

April 10, 2017

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Marc-André Gosselin**

Entitled: Understanding the influence basic service and value-added service usage has on customer churn in Pay-TV subscription services

and submitted in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE IN ADMINISTRATION
OPTION MARKETING

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

______________________ Chair
Dr. Charles Draimin

______________________ Examiner
Dr. Tieshan Li

______________________ Examiner
Dr. Jooseop Lim

______________________ Thesis Supervisor
Dr. Mrugank Thakor

Approved by: ____________________________________
Graduate Program Director

____________2017 ______________________________
Dean of School

ABSTRACT

Understanding the influence basic service and value-added service usage has on customer churn in Pay-TV subscription services

Marc-André Gosselin

Similarly to other North-American markets, new digital services and alternatives to the traditional pay-TV service are proliferating while the Canadian pay-TV industry is witnessing persistent subscriber losses. In an attempt to support changing viewing behaviors, generate more value and protect the subscriber base, pay-TV operators are extending their core TV service using value-added services (VAS). However, whether or not VAS successfully contributes to reducing subscriber attrition is unknown for academics and operators alike. Using survival analysis, the research examines VAS usage and churn behaviors for 11 647 pay-TV customers over a 12-month period. The results show that VAS users are not systematically less likely to churn and their churn behavior largely depends on usage frequency and usage patterns. Customers with constant or increasing usage frequency are less likely to churn than non VAS users and heaviest users appear to exhibit the greatest level of risk. Results also show that beneficial effects of VAS are generated by free services while payable VAS actually increases customers' risk. These findings show that churn prediction models need to look beyond the core service and examine actual behavioral usage statistics for both the core service and value-added services. From a managerial perspective, the results confirm that service extensions do indeed generate value and operators can further reduce customer attrition by maximizing VAS adoption. However, the results also show operators need to maintain and stimulate usage to preserve the beneficial effect of VAS and better understand the drivers that increase service switching behaviors.

ACKNOWLEDGEMENTS AND DEDICATION

Firstly, I would like to extend my gratitude for the guidance, patience and continuous feedback Professor Thakor has provided to me throughout this endeavor. As a part-time student balancing my academic and professional ambitions, Professor Thakor's continued interest and support was paramount to the success and completion of this Master's Thesis.

In addition, I am very grateful for Professor Li and Professor Lim's participation as members of my Thesis committee. Professor Li and Professor Lim have provided valuable and constructive feedback to maximize my contributions in preparation to the Thesis defense and final submission.

Lastly, I would like to express my deepest appreciation to my wife for her patience and her continued support for my academic and professional aspirations. I dedicate the efforts, persistence and sacrifices required to complete this Master's Thesis to the success and prosperity of my growing family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

In light of fierce competitive dynamics and new market trends, the pay-TV industry in Canada is losing terrain and the overall market share is under substantial pressure. In fact, the number of households that have permanently terminated their pay-TV subscription has gone from 193k in 2012 to 333k in 2015, and is forecasted to lose another 334k households in 2016 (The Convergence Consulting Group Limited, 2015). Luckily, traditional pay-TV operators have been able to offset losses by increasing ARPU, which generated a 1.7% increase in overall revenues for the 2015 calendar year (Advanced Television, 2016). However, given that technological advances are breaking down entry barriers and the traditionally closed pay-TV industry is witnessing the entry of new competing service alternatives, the compensation effect from active subscribers is unsustainable. These new service alternatives (i.e. Netflix, YouTube, Amazon, AppleTV, etc.) will likely continue to inflict significant pressure on the traditional subscriber base and even decrease subscriber acquisitions as younger consumers are likely more capable and willing to seek alternatives to the more expensive pay-TV subscription model. In fact, even if the pay-TV market share is under pressure, overall media and video consumption is actually on the rise across all age groups (Nielson, 2016). In addition, traditional media consumption including live TV and time-shifted TV using a DVR (i.e. digital video recorder) has remained constant while new media delivery over the Internet is consistently on the rise (Nielson, 2016). Therefore, it appears that customers are beginning to shift to new delivery methods, which may potentially accelerate the adoption of new competing services and fuel traditional pay-TV subscriber churn.

Zineldin (2006) emphasizes that companies need to generate customer relationships and experiences that create and deliver value beyond what is provided by the core product. In the hopes of creating more value, reducing customer churn, and adapting to new viewing behaviors, pay-TV operators are investing significant resources to enhance and extend their core service through various types of value-added services. For pay-TV operators, value-added services are characterized as "the availability of additional services over and above basic TV programs" (Lin et al. 2012). For example, pay-TV operators offer services like video-on-demand, ultra-high definition quality, and have recently expanded the core service to Internet connected devices (i.e. computers, mobile phones, tablets, etc.) to increase accessibility and enable new usage behaviors. Value-added services are recognized to drive growth through new revenues, act as a key differentiator for the operator (Kuo & Chen, 2006) and increase customer satisfaction (Jarrai, 2014). In addition to creating growth, value-added services are also recognized for their ability to generate a positive influence on a customer's intention to adopt digital-TV services (Ko et el., 2013; Madden et al. 1999). However, the role value-added services play in retaining customers is ambiguous.

Moreover, very few studies have explored the influence value-added services have on a customer's likelihood to churn. Although pay-TV and telecommunications literature boasts a variety of research models that have examined customer churn (Ku et al., 2011; Santouridis & Trivellas, 2010; Portela, 2010), details pertaining to value-added services are often overlooked. In addition, researchers have not made a clear distinction between the core service and value-added service and have primarily relied on customer expenditure to measure service usage. Kim and Yoon (2006) have attempted to examine the influence between the satisfaction towards value-added services and churn, but could not generate significant findings. In addition, Madden et al. (1999) examined whether or not the availability of value-added services influenced customers' likelihood to churn and also yielded insignificant results. According to Portela (2010), the most important predictor to influence churn in telecommunications concerns the amount a customer spends with the service provider and according to the researcher, actual service usage does not influence a customer's subscription duration. Given that pay-TV subscriber losses are currently offset by increases in ARPU, these findings suggest that pay-TV operators may witness more churn as monthly subscription costs increase. In addition, although Portela (2010) does suggest that usage does not seem to influence customer churn, usage behaviors with regard to value-added services have never been examined. In support to Portela (2010), Geetha and Kumari (2012) have found that customers that reach a certain threshold of value-added service expenditure are more susceptible to churn. However, this result can only be considered for services that increase customer spending and cannot be interpreted or even generalized across all types of value-added services, especially those that do not generate a direct cost to the customer. Therefore, while excessive payable value-added services do influence churn, a broader portrait regarding different types of value-added services and customers' underlying usage behavior needs to be examined.

Therefore, the primary objective and contribution of the research is to look beyond the core service and truly understand the different effects value-added services have on customer churn with pay-TV services. In addition, because a customer's perceived value with regard to a product or service is the combination of the value delivered by the core service and the value delivered by value-added services (Gronroos, 2004), researchers examining customer churn on the basis of the core service are excluding an important component to churn analysis. In addition, even if Portela (2010) does not recognize usage as a significant predictor of churn, Li et al. (2015) argue that actual behavioral variables, including usage, will generate more accurate churn predictions than solely relying on subscription, billing, and demographic variables. Therefore, the research will attempt to answer the two following questions:

1) *Are users of VAS (value-added services) less likely to churn than non-users of VAS?*
2) *How do different VAS usage behaviors influence a customer's likelihood to churn?*

From a theoretical standpoint, the results of this research will contribute to existing churn literature by providing a broader outlook to components that influence churn in pay-TV subscription services. That is, the research will go beyond basic subscription and demographic predictors and examine how service and more specifically, value-added service usage may contribute to customer survival or better predict customer defection. In addition, the research will also provide greater granularity on how different types of value-added services may have different effects on customers' subscription duration. These results will contribute to existing literature by providing more insight to the role value-added services play in churn prediction which will hopefully generate a new direction for research and supplement current churn prediction models. From a managerial perspective, operators have invested and are continuing to invest significant resources to develop, improve, operate, and support various value-added services to generate more value for customers and respond to new competitive service alternatives. However, whether or not these investments are having beneficial effects on the customer base is ambiguous for operators. That is, pay-TV operators can expect to attract new customers and generate new revenues, but cannot ascertain the true potential value-added services have on existing customer relationships. Therefore, this research will provide pay-TV operators better comprehension on the influence value-added services have on the existing subscriber base, whether or not usage patterns are indicative of a customer's churn or survival behavior and whether or not their effort to enhance the core service does indeed contribute to customer loyalty. These results will also provide pay-TV operators additional insight to identify and target high-risk customers for proactive retention campaigns.

LITTERATURE REVIEW

A substantial number of researchers and studies have examined various factors that influence subscriber churn in subscription service contexts, and especially with telecommunication services. Firstly, for telecommunication operators, a churned subscriber is defined as a subscriber that, for any given month, has paid the service fee for the previous month and stopped paying the fee for the current month; the time at which the subscriber cancels the service subscription (Li et al., 2015). Although, churn analysis is prominent in telecommunications, many of these studies concern mobile and Internet services and few have explicitly explored customer attrition with regards to pay-TV services. In addition to studies specific to pay-TV services (Burez & Poel, 2007; Jamal & Bucklin, 2006; Ko et al. 2013; Lin et al. 2012; Prince & Greenstein, 2014), certain results from other churn analyses for mobile and Internet services can also be generalized to the pay-TV service. However, among those results, there are factors that cannot be as easily generalized across all types of services because the nature of the business model varies from one service to another. For instance, the pay-TV business model is primarily characterized by a fixed monthly subscription fee while usage limits and surcharges for excess usage are much more likely for mobile and Internet services. Therefore, even if customer expenditure has been recognized as an accurate predictor for usage (Geetha & Kumari, 2011), this measure cannot be generalized to pay-TV services because the nature of the business model is inherently different. That being said, the way customer expenditure and usage are associated is different for pay-TV services, and consequently, the two predictors are also likely to behave and influence churn differently than other telecommunications services. Furthermore, many of these studies and churn prediction models solely rely on general and static billing information, fail to capture behavioral variables and most importantly, overlook the difference between basic and value-added service usage.

General predictors of customer churn

Factors that influence customer churn in the telecommunications industry and broader contractual and subscription services include customer satisfaction (Bolton, 1998; Ranaweera & Prabhu, 2002; Svendsen & Prebensen, 2011), trust customers have towards the service provider (Ranaweera & Prabhu, 2002), brand image (Svendsen & Prebensen, 2011), customer demographics, perceived switching cost, switching barriers (Svendsen & Prebensen, 2011; Ranaweera & Prabhu, 2002), knowledge and awareness of alternatives (Caparo et al. 2003). While these factors are accurate and generally recognized predictors of customer churn, they are not used in churn prediction models because this data is not necessarily readily available in companies' customer relationship management systems and do not capture actual customer behavior. Also, this cross-sectional data makes it more tedious and time consuming for companies to systematically and continually measure customer behavior and customers' likelihood to churn on an ongoing basis. This limitation also limits a company's predictive capability to identify and target high risk customer in real-time and implement targeted retention programs. Consequently, many churn analyses in telecommunications rely on readily available customer data to help managers use existing information more efficiently and build prediction models using data from billing and CRM (i.e. customer relationship management) databases (Jamal & Bucklin, 2006). These prediction models rely on *subscription variables* such as monthly service expenditure (Anh et al. 2006; Burez & Poel, 2007; Portela et al. 2010), contract duration and time to expiration, the options a customer has (i.e. sports, movies…), the type of digital set-top box, (Burez & Poel, 2007), other products/services a customer has with the operator and whether or not these services are bundled to other services (i.e. Internet, telephone, mobility…) (Jamal & Bucklin, 2006; Burez & Poel, 2007; Prince & Greenstein, 2014); *socio-demographic variables* such as age, gender and a customer's region (Burez & Poel, 2007); *financial variables* like customer debt, overdue bills, billing reminders and even customers' payment method (Anh et al. 2006; Burez & Poel, 2007), and lastly; *interaction variables* which include customer complaints (Anh et al., 2006), number of service interactions for service related issues (i.e. billing, service modifications, customer support …) and service recovery attempts (Jamal & Bucklin, 2006). As previously suggested, these variables do not necessarily capture behavioral changes that may foreshadow churning behavior and Li et al. (2015) demonstrated that churn prediction models that include behavioral data such as usage, usage frequency, and usage patterns could significantly improve prediction performance.

Moreover, to predict customer behavior using behavioral and attitudinal variables without systematically collecting cross-sectional data, Zorn et al. (2010) rely on proxy indicators for satisfaction and customer commitment, each of which are recognized as important components to understanding churn behavior. According to this study, Zorn et al. (2010) have demonstrated that the occurrence with which a customer changes its subscription plan is in fact an indicator of customer commitment. Zorn et al. (2010) also show that the more a customer changes its subscription plan, the customer becomes less committed and more likely to quit the relationship. Although service usage frequency is different than a change in the subscription plan, altering usage patterns may also indirectly capture customers' level of interest and commitment with the service. Jamal and Bucklin (2006) also used proxy variables to define broader qualitative indicators of churn and include: "failure recovery" (i.e. a customer's satisfaction with a recovery attempt), "payment equity" (i.e. a customer's evaluation of the cost-benefit trade-off and satisfaction), and "customer service experience" (i.e. a customer's evaluation of the provider's customer service efforts). Through their model, Jamal and Bucklin (2006) show that there are different types of churners; active churners are customer that communicate with the firm about their problems while passive churners have low contact rates and churn without communicating with the firm. These results suggest that behavioral predictive measures are important to adequately target and capture passive churners. Even if this study will not rely on proxy variables to qualify churn predictors and churn behavior, these findings show that the study's interest in behavioral service usage statistics is justified and an essential component to adequately measure churn behavior.

Furthermore, even if customer expenditure does not necessarily reflect customer behavior, Portela (2010) argues that it is the most important predictor of churn and usage does not have a significant effect. On the other hand, Anh et al. (2006) suggest that usage is one of the most popular predictors of a customer's behavioral intentions to churn. In fact, there is ample empirical evidence to show that highly satisfied customers have longer lasting relationships with a service provider (Bolton 1998) and also have higher usage levels (Bolton & Lemon 1998). In fact, several studies with telecommunication services suggest that service usage, measured by monthly expenditure, is positively associated to churn probability and that heavy users are most likely to churn (Madden et al., 1999; Anh et al., 2006). Burez and Poel (2007) also measure usage using financial indicators because of the broadcasting nature and the technological limitations to colleting actual viewing statistics. However, when examining actual usage behavior for mobile services (ex. minutes of usage, number of calls, number of short messages - *SMS*), Kermati and Ardabili (2011) found that usage is negatively associated with churn probability. These findings are more consistent with those of Bolton (1998) and Bolton and Lemon (1998). This divergence can attributed to the fact that usage in churn analysis is most commonly measured using financial indicators rather than actual service usage statistics. Although Geetha and Kumari (2011) suggest that both measures are good predictors of customer behavior, Li et al. (2015) have demonstrated that models that account for actual behavioral usage will better predict customer churn events than static billing information. In addition, while financial measures capture transactional usage accounted for in customers' monthly spending, this measure will fail to capture usage for services that do not generate a cost to the customer. Given that monthly service fees for pay-TV are commonly structured on fixed subscription models through channel and service selection, customer spending is not expected to significantly vary from one observation period to another. Although variations in customer spending can be generated by transactional purchases and rentals, this expenditure is typically a small proportion of a customer's monthly service fee and this measure will therefore overlook a significant proportion of service usage. Therefore, in light of these findings and contradiction, it is expected that pay-TV service usage frequency will be negatively related to churn. Furthermore, Reinartz and Kumar (2003) argue that if a customer devotes a large "share-of-wallet" to a company, the customer relationship will be stronger and last longer. Although this study is not as interested in customer spending than service usage, a parallel can be inferred to "share-of-use". That is, customers attributing a greater portion of their entertainment viewing time to the pay-TV service rather than other service substitutes are expected to display longer lasting relationships with the pay-TV service provider. Therefore, it is anticipated that greater service usage will be positively associated to the customer's relationship duration with the service provider.

In addition, current studies in telecommunications, including pay-TV services, capture customer expenditure or usage for the overall service and do not make the distinction between usage generated by basic and value-added services. Therefore, to isolate the effects of basic service usage and value-added service usage, it is hypothesized that basic service usage (in viewing hours) will have a negative influence on customer churn.

H1: *Basic (core) service usage has a negative influence on customer churn.*

Value-added services and the pay-TV industry

Value-added services (VAS) are characterized as “the availability of additional services over and above basic TV programs” and may include services such as video-on-demand, a personal video-recorder, premium channels, high-definition and ultra-high definition, 3D services, etc. (Lin et al. 2012; Ko et al., 2013). To provide a greater and a more convenient access to television content (i.e. subscription content, linear channels and video-on-demand), pay-TV operators have also extended the basic TV service using pay-TV applications accessible to Internet-connected devices (i.e. smart-phones, tablets, laptops, computers, internet-connected TVs, etc.) and are typically referred to as online-TV applications. Some of the above mentioned services are provided at no additional cost to customers (free VAS) and are intended to provide more value to the basic service and pay-TV subscription while others are transactional services above the basic pay-TV subscription (payable VAS). While payable VAS is also recognized to drive customer value and fulfill additional customer needs, they also generate additional revenue and growth for service operators.

Ko et al. (2013) categorize the pay-TV service into three specific dimensions: the basic service, value-added services, and interactive services. Specifically, the *basic product / service* is defined as “basic and necessary product / service functions that may cater to DTV (digital TV) viewer’s needs in terms of audio-visual effect and operational functions of DTV”, *value-added services* as “digital multi-channel combination and displays that may enhance DTV (digital TV) viewers’ general value-added service functions on the basis of necessary functions”, and lastly, *interactive services* as “a process that may enhance DTV viewers’ participation and emotions by means of increasing interactivity on the bases of both the basic product / service and general value-added functions” (Ko et al., 2013). Therefore, the basic pay-TV service includes access to traditional linear channels and broadcasting using the digital set-top box, value-added services as any extension to the basic service designed to enhance the basic functions of the core service such as video-on-demand services, digital video-recorder, online-TV applications, and lastly, interactive services as the ability to interact with the service through features that enable interactions on social media and interactive gaming.

Value-added services as a predictor of customer churn

Krishnan and Kothari (2006) found that the best indicators to determine whether or not individuals will recommend a service provider are customer service and the variety of services offered (i.e. value-added services). In addition, according to Oliver and Winer's (1987) utility framework, customers that buy more, buy more frequently and across a variety of categories derive greater utility, have better fit with the vendor, and consequently, have longer lasting customer relationships. Reinartz and Kumar (2003) validated this concept by demonstrating that customers with a broader scope of interactions with a vendor through cross buying have longer lasting and stronger relationships with that vendor. Although this utility framework is interpreted through product purchasing contexts rather than a service context, customers that adopt and use a larger breadth of services (i.e. value-added services) also have a greater number of interactions with the provider's service offering which may generate similar effects on customers' relationship duration. In fact, in telecommunications and pay-TV services, researchers have found that value-added services are an important component that influences a customer's likelihood to adopt a new service (Ko et al., 2013; Ku et al., 2009; Madden et al., 2009). These findings suggest that VAS generates value for customers, and according to Wang et al. (2004), customer value has a significant effect on customer satisfaction, customer retention and long-term profitability (Ku et al. 2011). These finding also suggest that Oliver and Winer's (1987) utility framework may be interpreted and applied to telecommunication service contexts. Although Madden et al. (2009) could not establish a significant relationship between the availability of value-added services from ISPs (*Internet Service Providers)* and customers' churn probability, the findings did establish a positive relationship between their availability and customer satisfaction. In fact, numerous studies have demonstrated that VAS has a positive effect on customer satisfaction (Jasrai, 2006; Kim et al., 2004; Lam et al., 2004; Lim et al., 2006; Unhanandana & Wattanasupachoke, 2012; Zhang et al., 2014) and service quality (Choi et al., 2007; Santouridis & Trivellas, 2010; Gerpott et al., 2001; Kim et al., 2004), each of which have been recognized as some of the most important predictors of customer churn in telecommunications. However, these results only stipulate that VAS may have an indirect influence on customer churn.

Numerous studies have proposed models to measure the relationship between VAS, service quality (Kim & Yoon 2004; Ku et al., 2011) and loyalty (Madden et al., 2009; Santouridis & Trivellas, 2010), but each have failed to yield significant relationships. However, even if the study has not examined customer outcomes with regard to the actual churn event, Unhanandana and Wattanasupachoke (2012) have found that promotional advantages and VAS generate positive attitude and customer loyalty. These findings corroborate with Lin et al. (2012) that have found that VAS contributes to a service's perceived benefit and customers' intention to remain subscribed to the pay-TV service. However, these findings only stipulate that VAS has an influence on a customers' intention to remain subscribed to the service and does not provide insight to the actual relationship between VAS usage and the customer outcome (i.e. the churn event or sustained survivability). Therefore, by examining actual VAS usage and customer churn, it is hypothesized that sustained VAS users will be less likely to churn than non-users.

H2a: *VAS users are less likely to churn than non-users.*

Although users of VAS are hypothesized to display lower risks of canceling their subscription, it is anticipated that the different types of VAS (i.e. free vs. paid) have a different influence on this risk. Similarly to customer expenditure, there is evidence that excessive payable VAS usage with mobile services will increase customers' likelihood to churn. In fact, Geetha and Kumari (2011) have found that customers with payable VAS expenditure reaching over 30% of the overall monthly service fee were more likely to churn. The study argues that excessive VAS usage indicates that the benefits derived from the core service is less and consequently, these customers were more susceptible to churn (Geetha & Kumari, 2011). However, these results are only applicable for the heavyset users of payable VAS and do not provide any insight with regards to the different types of VAS usage or different frequencies of use. That is, these finding do not account for the influence of free VAS or lighter and medium usage frequency. Furthermore, Unhanandana and Wattanasupachoke (2012) argue that the influence promotion advantages and VAS have on customer loyalty occurs because "customers pay great attention to marketing strategies such as special discounts, award credits, free premium products, bonus rewards, and free product/services with the regular purchase". Although free VAS in pay-TV is not offered through discounts or credits, they are offered as a free extension to the basic TV service subscription. These findings suggest that free VAS and payable VAS do not generate the same value or perceived benefit for customers. In fact, the most common conceptualization of value has been in evaluating the trade-off between the benefits (i.e. what the customer gets) and the sacrifices (what a customer gives) regarding the attributes of a product or service (Sanchez-Fernandez & Iniesta-Bonillo, 2007; Zeithaml, 1988). Precisely, Zeithaml (1988) argues that value is defined by "the quality I get for the price I pay". This definition implies that customers will derive greater utility and value when the tradeoff between what is given and received is minimized. While there is a dependency on the utility derived from the service, it is reasonable to assume that customer spending may reduce a customer's utility and perception of value. This concept is consistent with Jamal and Bucklin's (2006) findings that a customer's evaluation of the cost-benefit tradeoff (i.e. payment equity) influences satisfaction and churn propensity. Through the conceptualization of payment equity, Jamal and Bucklin (2006) argue; "as the cumulative amount invoiced increases, we expect cumulative benefits to outweigh the cumulative cost, leading to a decline in churn rates". Furthermore, Jamal and Bucklin (2006) also suggest that diminishing marginal returns from subscribing to additional services and increasing expenditure is offset by the utility derived by the higher value package. Thus, if this value offsets the effects additional spending has on churn likelihood, value generated by free value-added services are likely to further increase a customer's evaluation of the cost-benefit tradeoff and consequently, have a greater influence on reducing churn likelihood than payable value-added services. Therefore, the different types of value-added services (free vs. paid) are expected to yield different influences on customer churn. It is therefore hypothesized that free VAS usage will have a greater influence on reducing churn susceptibility than payable VAS usage.

H2b: *Free VAS usage has a greater influence on reducing churn susceptibility than payable VAS* usage.

Furthermore, VAS users with different usage frequencies and usage patterns are also expected to display different churn behaviors. In fact, Reinartz and Kumar (2003) argue that customers with greater interaction frequencies with a vendor also have longer lasting relationships. Therefore, a greater interaction frequency with the service (i.e. usage) is expected to be associated with longer lasting subscription durations. Although not specific to VAS, Keaveney and Parthasarathy (2001) argue that heavier users of online services are less likely to switch because of the disconfirmation paradigm which specifies that "frequent usage should provide customers with relatively accurate and realistic performance expectations, thereby decreasing disconfirmation and increasing satisfaction and repurchase intentions". In fact, Keaveney and Parthasarathy (2001) found that service continuers had greater usage and more prior experiences with the service than did switchers. The researchers argue that customers that frequently use the service develop strong and positive attitudes towards it, thereby increasing satisfaction (Keaveney & Parthasarathy, 2001). In addition, according to Lee et al. (2001), users with the highest usage frequency (i.e. heavy-users) also have strong attachments to VAS, which also increases their satisfaction. Although Geetha and Kumari (2011) found that the heaviest users of VAS lead to greater churn, their results are only true for excessive use of payable services and cannot be generalized to overall usage of VAS. In addition, Geetha and Kumari (2011) only looked at the effect payable VAS had above a specific threshold and did not study the effects of VAS below that threshold. Also, Geetha and Kumari (2011) examined VAS spending rather than VAS usage behaviors and patterns thereby supporting similar findings that suggest increased customer spending increases churn susceptibility. Therefore, given that this research is interested in understanding how behavioral usage patterns affect churn likelihood, there is ample evidence to suggest that increased VAS usage frequency among VAS users will further decrease churn susceptibility. Therefore, it is hypothesized that users with greater overall VAS usage frequency (paid + free) will be less likely to churn than those with lower usage frequency.

> ***H3a:*** *VAS users with greater overall (paid + free) VAS usage frequency are less likely to churn than those with lower usage frequency.*

In addition, Anh et al. (2006) suggest that customers do not suddenly churn and switch to a new service provider. Rather, a customer's usage status may change prior to the cancellation event and may in fact provide evidence that a customer is potentially going to churn. In fact, Anh et al. (2006) show that customers who temporarily suspend their account and change to an inactive user status become more likely to churn. In addition, given that increasing usage frequency is recognized among service continuers (Keaveney & Parthasarathy, 2001) a change in a customer's VAS usage pattern (i.e. constant, increasing, decreasing) may also be evidence that a customer is becoming more or less likely to churn. In fact, Allenby et al. (1999) have developed a model to recognize when customers change their purchasing patterns and show signs of defection. Precisely, Allenby et al.'s (1999) model identifies and qualifies a customer's status from a "super-active, active and non-active" state and when customer moves from one state to another. While this model was initially developed to better plan direct marketing communications, knowing when a customer moves from a "super-active", to "active", or "non-active" state is very valuable to identifying customer's that are more likely to churn. Inspired by this model, a change in a customer's VAS usage state (i.e. decreasing from a medium to light user) may be an indication that a customer exhibits greater churn risk while a constant or increasing VAS usage state may be an indication that customer churn is less likely. Furthermore, changes in customer behavior such as service upgrades or downgrades may also reflect a change a customer's interest and commitment to the service (Zorn et al., 2010). In fact, Zorn et al. (2010) have qualified customer commitment by measuring the number of times customers change their subscription plan and this measure has been recognized as an important attitudinal predictor of churn. Although behavioral usage patterns (i.e. constant, increasing, decreasing) is a different measure than subscription plan changes, sudden decreases in VAS usage may also indicate that a customer is losing interest, becoming less committed to the service and more likely to terminate the service. Thus, it is hypothesized that VAS users with increasing VAS usage pattern are less likely to churn than VAS users with a decreasing usage pattern.

H3b: *VAS users with increasing VAS usage pattern are less likely to churn than VAS users with decreasing usage pattern.*

Furthermore, inspired by customer learning theory, Keaveney and Parthasarathy (2001) suggest that heavy users of online services may be less likely to switch service providers because of nontransferable provider specific skills acquired through increased usage. The theory also suggests that customers with the acquired skills and knowledge may be unwilling to learn how to use alternative products and services (Keaveney & Parthasarathy, 2001). However, given the accessibility of alternatives to traditional pay-TV services, heaviest users may also be more informed and aware of these substitutes and consequently, more likely to switch to competing services. Even if acquired skills may be a deterrent to switching services, Anh et al. (2006) argue that heavy users with accumulated service experience may in fact be more likely to explore new more advanced services. In fact, customer experience has been recognized to be both beneficial and detrimental to behavioral loyalty (Dover & Merthi, 2006). That is, while greater and increasing usage frequency is expected to have beneficial effects on churn, heavy users that have accumulated experience and knowledge are also more aware of service substitutes which makes it easier for these users to change to competing offerings (Zorn et al. 2010). These findings are consistent with Caparo et al.'s (2003) study revealing that customers' knowledge and awareness of substitutes increased churn likelihood. In addition, highly competitive markets have greater targeted marketing communications and advertising tactics making customers even more aware and knowledgeable (Bolton et al. 2004). Therefore, given the highly competitive nature of the pay-TV industry and the increasing amount of new online substitutes, customers aware of these substitutes may be attracted by their novelty and more likely to switch to these new offerings. Thus, even if pay-TV service providers are attempting to innovate and are extending the basic service with value-added services, heaviest users will be more aware of service substitutes than any other user-group and are also expected to exhibit the greatest level and risk of churn, even in comparison to the control group of non-users. Therefore, beneficial effects of VAS usage frequency and usage status (i.e. increasing) is expected to reach a certain threshold where heaviest users actually become more likely to churn to explore new alternatives.

H3c: *Heaviest VAS users exhibit the greatest risk of churning.*

Therefore, in addition to examining and validating generally recognized predictors of customer churn specific to the pay-TV industry, this study will also provide new behavioral churn predictors by measuring and deconstructing how pay-TV VAS usage affects customer churn and determining whether or not different types of services (i.e. free vs. paid) influence churn differently. Although Madden et al. (2009) and Santouridis and Trivellas (2010) have failed to yield significant findings for VAS, these studies relied on cross-sectional data, which may not adequately capture true behavior and churn likelihood.

Researchers have adopted various different approaches to study customer attrition and build customer churn prediction models in telecommunications. For instance, logistic regression is commonly used to measure the influence predictors have on a customer's churn likelihood and is recognized for its ability to successfully predict churn (Lu, 2002). However, logistical regression does have a major drawback because it is unable to measure how a customer's risk changes over time and assumes that this risk is constant (Van den Poel & Larivière, 2004). On the other hand, survival analysis and hazard models have become recognized as a robust method to analyze duration data and yield more accurate results than traditional methods (i.e. logistic regression, least squares regression, decision trees, etc.) (Lu, 2002; Jamal & Bucklin, 2006; Van den Poel & Larivière, 2004). In fact, survival analysis yields superior results than traditional methods because of its ability to better support duration data, integrate variables that change over time (i.e. time dependent covariates), and allow for a customer's risk to change over time (Lu, 2002; Jamal & Bucklin, 2006). In addition, for datasets with fewer churn events, survival analysis captures much more information than logistic regression because the model has the ability to interpret information from subjects that have not generated a churn event. That is, contrary to survival analysis, logistic regression considers a "0" or "1" churn outcome and ignores the effects of a customer's survival time when the cancellation event does not occur (Kleinbaum & Klein, 2012). Therefore, the study will rely on survival analysis to better predict customer churn and better understand how time-varying covariates affect customers' likelihood to churn.

Survival analysis and hazard models

Survival analysis is composed of a set of statistical tools and methods to study the occurrence and timing of an event using longitudinal data. Survival analysis is defined as "the time to the occurrence of a predefined terminating event" (Lai & Zeng, 2013). In the context of this study, the event of interest occurs when a customer terminates the subscription (i.e. churn). The origins of survival analysis can be traced back to medical studies and have since been used across multiple different disciplines including marketing and customer attrition analysis. By studying the risk to an event, survival analysis enables researchers to estimate and interpret customers' churn risk, compare churn likelihood between two or more groups, and assess the relationship between explanatory variables and customers' risk to churn (Kleinbaum & Klein, 2012). Within this study, a customer's survival time is the time in months the customer will remain subscribed to the pay-TV service, and the "event" or "failure" occurs when a customer terminates the pay-TV subscription. In survival analysis, $\mathrm{T} = \text{survival time } (\mathrm{T} \geq 0)$, where T is a random variable with its own probability distribution. Another attribute where survival analysis outperforms conventional approaches is its ability to manage censored data (Lu, 2002). Precisely, censoring occurs when the dataset includes information on customers' survival time without knowing the exact survival time. Although the subject may not have churned during the observation period and what happens following the observation period is unknown, the subject may nonetheless have a greater risk of churning. The most common type of censorship is "right-censored" which means; "true survival time is equal or greater than observed survival time" (Kleinbaum & Klein, 2012). That is, censored observations are from research subjects for which the event in question did not occur by the end of the study, but may nonetheless occur outside of the research period. Therefore, even if the event did not occur, survival analysis can provide insight on which factors may contribute to a customer's risk of churning in the future.

The most import components of survival analysis are the survival function, the hazard function, and the hazard ratio. The survival function and hazard function define customers' survival status and instantaneous risk while the hazard ratio is used to measure the effects covariates have on the hazard function. The survival function $S(t) = P(T \geq t)$ gives the probability that a customer will survive longer than the specified period of time denoted by t (Kleinbaum & Klein, 2012). On the other hand, the hazard function $h(t)$ gives "the instantaneous potential per unit time for the event to occur, given that the individual has survived to time *(t)* (Kleinbaum & Klein, 2012). Therefore, the survival function focuses on a customer's survival time while the hazard function focuses on the failure event. Unlike the survival function, the hazard function does not yield a probability, but rather the potential that the event will occur for each unit of time. As denoted in the notation below, the hazard function includes a probability statement divided by a change in time, which will yield a probability per unit of time. Therefore, the scale for this ratio is not that of a probability (0 to 1), but will range between 0 and infinity depending on the unit of time that is used (Kleinbaum & Klein, 2012).

$$\lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}$$

Lastly, in survival analysis, the measure effects variables have on the hazard rate is referred to as the hazard ratio. In fact, similarly to logistic regression (i.e. odds ratio), the hazard ratio for a covariate is $e^{\beta}$and represents the relationship between the covariate in question and the survival time T (Kleinbaum & Klein, 2012). For a continuous variable, the hazard ratio is "the ratio of hazards for every unit increase in the predictor variable" while for a categorical predictor, the hazard ratio is "the ratio of the hazard rates between the given category and the reference category" (SAS® Institute Inc., 2016).

There are several different methods and statistical models to conduct survival analysis. This first model is a nonparametric model that is typically used as a first step to estimate the sample population's survival function or compare survival functions between two or more groups (Gardiner, 2010). The model is referred to as nonparametric because there are no underlying assumptions regarding the hazard function's functional form. This is achieved by estimating the survival function of the sample population or sub-groups by using the Kaplan Maier estimator, also known as the product limit estimator (Kaplan & Meier, 1958). Kaplan Meier curves will plot and depict the sample population's survival function and will enable researchers to compare the different survival functions for each sub-group of interest.

In contrast, parametric models assume that survival functions are not constant over time and follow a pre-determined pattern. The form and survival distribution can take various different shapes (i.e. Weibull, exponential, log-logistic, log-normal, and gamma) and that shape may depend on time, a set of covariates, or both (Kauffman & Wang, 2008). For example, in the medical field, a decreasing Weibull distribution would suggest that patients that were operated have a decreasing risk of dying as the time after the surgery increases. The opposite may be true for subscription services as a customer's risk of canceling may increase as a function of time (i.e. increasing Weibull distribution). In fact, Portela (2010) found that churn distribution in telecommunications is neither constant over time nor across customers. Within this context, full parametric models provide more accurate results as the survival and hazards functions have a better fit with the survival data because the standard error within the model is reduced (Lu, 2002). However, this requires researchers to predict and pre-determine the exact shape of the population's survival function. Pre-determining the shape of this distribution is not always possible nor is it trivial as it goes beyond examining the shape of the sample's survival data. That is, because many of the observations within a sample dataset may be censored, the occurrence of the event happening outside of the observation period may very well influence the population's survival distribution and any assumption made with regard to the sample's distribution may be inaccurate or false. In addition, although paramedic models may provide more accurate results, they do not handle time-dependent covariates, which is a significant drawback compared to other survival analysis methods (Liu, 2010). In fact, Van den Poel and Larivière (2003) advocate for models to incorporate variables with different values over time because they increase model performance (Van den Poel & Leunis, 1998) and yield more accurate forecasts (Weerahandi & Moitra, 1995).

The Cox Proportional Hazard (PH) model is a semi-parametric model with the ability to handle time-dependent covariates and has few restrictive assumptions. In fact, the Cox proportional hazard model dominates in the field of dynamic survival models across several different disciplines and is recognized as a very robust model (Van den Poel & Larivière, 2003). In addition, the PH model will even approximate the results of a correctly fitted parametric model (Kleinbaum & Klein, 2012). The Cox PH model is interpreted by $h(t,X) = h_o(t)e^{\sum_{i=1}^{p_1} \beta_i X_i}$ where $h_o(t)$ is the baseline hazard function involving time *(t)*, and $e^{\sum_{i-1}^{P} \beta_i X_i}$ computes the hazard ratio for each time independent explanatory variable X_i with coefficient β_i (Kleinbaum & Klein, 2012). Although the approach does not impose restrictions on the survival distribution, the model does assume proportional hazards, which means "the hazard for one individual is proportional to the hazard of any other individual, where the proportionality constant is independent of time" (Kleinbaum & Klein, 2012). This assumption also applies to the influence covariates have on the hazard rate meaning the rate of hazard is constant thereby avoiding temporal biases to become influential. However, this assumption only applies to covariates that are fixed in time and time-dependent covariates are an exception to this assumption (Poel & Larivière, 2004). That is, because time-dependent variables will take different values over time, this time may impose varying effects on the rate of hazards. When the PH assumption is not met for time-dependent variables, an extended version of the model can adjust and incorporate the effects for time-dependencies. The extended model is interpreted by $h(t,X(t)) = h_o(t)\left[\sum_{i=1}^{p_1}\beta_i X_i + \sum_{j=1}^{p_2}\delta_j X_j\right]$ where $X(t) = X_1, X_2, \ldots, X_{p1}$ are time-independent predictors and $X_1(t), X_2(t), \ldots, X_{p2}(t)$ are time-dependent predictors, and δ_j is the coefficient for $X_1(t)$ (Kleinbaum & Klein, 2012). That is, the extended Cox model adds time interactions to variables that vary with time and violates the proportional hazards assumption.

Dataset and data specifications

The dataset used for the study was built using real field data obtained from a major Canadian telecommunications operator. The operator provides a complete portfolio of subscription services for pay-TV, Internet, fixed home telephony and mobile services. Unlike other markets, the telecommunications industry in Canada is highly regulated by the *CRTC* (Canadian Radio-Television Communications Commission) and prohibits fixed term contracts for telecommunications services. Therefore, although customers commit to an ongoing monthly subscription, subscribers are free to disengage and cancel the subscription at any given time without prohibitive penalties. Therefore, the dataset and customers' survival distribution is not influenced by fixed-term contracts.

The dataset is composed of a random sample of the operator's pay-TV customer population and includes 13 195 anonymous residential pay-TV accounts. The dataset was built by combining various databases and includes customers' monthly *subscription variables* (ex. overall service fee, transactional purchases, premium services, type of digital set-top box, other bundled services, overdue bill amount, etc.), *demographic variables* (age, region, type of home) and monthly *usage metrics* (ex. overall usage, value-added service usage, free value-added service usage, and paid value-added service usage). Therefore, in addition to basic and value-added service usage metrics, the research will also consider the effects of control variables summarized in Table 1 (Appendix A), the majority of which are recognized for their effects on churn within the customer attrition literature.

Given that the research is interested in behavioral usage, basic service usage is defined and measured as the amount of time customers watch basic broadcasted content using the service provider's digital set-top box. Given that VAS are services that enhance the basic and necessary functions of the basic service (Ko et al., 2013), VAS is defined and measured as the amount of time spent watching content that goes beyond the basic linear broadcast (i.e. Video-on-Demand services or any time spent watching content on the operator's online-TV applications). Precisely, Video-on-Demand content includes; 1) channel-based Video-on-Demand - CVoD (i.e. previously broadcasted content made available after it has aired); 2) free Video-on-Demand - FVoD (i.e. any On-Demand content made available by the service operator for free); 3) payable / transactional purchases for ad-hoc video-rentals – TVoD; 4) subscription video-on-demand - SVoD (i.e. an unlimited access to a pre-determined catalogue of On-demand content given a monthly subscription). Given that the research is also interested in understanding the difference between free value-added services (free VAS) and payable value-added services (payable VAS), free VAS accounts for services that do not generate a direct cost to customers (i.e. CVoD, FVoD, online-TV) while paid VAS accounts for any type of usage generated by payable services (TVoD, SVoD).

Research Design

Similarly to other churn analyses in telecommunications literature and other service industries, the study will rely on a two-step survival analysis (Lu, 2002; Jamal & Bucklin, 2006). The first step of the survival analysis relies on the Kaplan Meier estimator to quantify initial survival functions and compare survival distributions between different user-groups. In fact, Prinzie and Van den Poel (2006) significantly improved their churn prediction accuracy in financial services by clustering customers according to the changes in their account balance over time. By adopting the same approach for VAS usage, this design is not only expected to demonstrate how usage patterns influence customer churn, it should also improve the model's prediction accuracy. Therefore, each user-group will reflect the different usage patterns found within the sample population and will discriminate between non-users, light-users, medium-users and heavy-users. The different survival functions depicted by the Kaplan-Meier curves show each of the group's survival distributions and should provide evidence on the influence VAS usage frequencies and patterns have on customer attrition. The second step of the survival analysis will be to build a prediction model using the Cox Proportional Hazard model. This model will provide a granular view on how the different explanatory variables influence customers' churn likelihood. The Kaplan Meier estimator will only provide insight on the different survival patterns between the user groups while the hazard model will scrutinize the relationship explanatory variables have on subjects' hazard ratio and ultimately, their likelihood to churn. This second step of the analysis will depict a much clearer notion on the predictors that contribute to a customer's survival or churn risk.

According to Lu (2002), the best observation approach in survival analysis is "prospective", meaning that the observation begins at a specific point in time (i.e. origin time) during which churn events are recorded for a substantial period of time until the end of the study (i.e. termination time). For this study, the longitudinal dataset includes a twelve-month period from September 2014 to August 2015. The dataset's monthly billing cycles begin on the 15th of every month, meaning that the study's origin time is September 15th 2014 and the termination time is August 14th 2015. The explanatory variables in the dataset can vary with time (i.e. time-dependent) on a monthly basis. For example, service usage metrics, customer expenditure and a customer's subscription duration are continuous and vary from one billing cycle to the next. In addition, categorical variables such as a customer's type of home, region and whether or not the set-top box is leased can also vary during the study period. The only time-independent variable in the model is the user-group subjects belong to (i.e. light, medium, heavy user). In addition, the dataset's monthly intervals are aligned with each customer's billing cycle, or "follow-ups" as referred to in survival analysis. For each subject, the dataset includes a unique and anonymous customer identification number, monthly follow-up for all explanatory variables in Table 1, an indicator to identify which user-group the subject belongs to, an indicator to identify whether or not the subject is censored, and lastly, the duration to the event whether the subject is censored (i.e. cancellation event did not happen) or not (i.e. the cancellation event occurred). The style of this input dataset is referred to as the Counting Process (CP) format and is useful to capture variations in time-dependent covariates (Kleinbaum & Klein, 2012). To best fit the model and measure account suspensions related to residential pay-TV services, the initial dataset was filtered to exclude commercial accounts, seasonal account suspensions, and accounts with missing data entries.

In addition, given that the study is interested in understanding how usage frequency and usage patterns influence a customer's likelihood to churn, usage observations are separated into two distinctive periods. This distinction is necessary to adequately measure whether or not increasing, constant or decreasing usage patterns have different effects on customers' outcome. Comparing these two periods and their underlying usage patterns will facilitate user-group segmentation and will rely on the three patterns of interest (i.e. constant, increasing or decreasing usage). Otherwise, given the size of the large sample size and the number of usage periods, the quantity of scenarios is likely to generate too many usage profiles thereby limiting successful clustering and interpretation. In addition, because the Canadian pay-TV service is highly dependent on a four-month seasonal programming calendar (e.g. fall / winter programming), grouping the usage periods according to the programming periods will also isolate seasonal usage fluctuations that may otherwise bias the results. Therefore, the research design was aligned to each programming period and was then followed by a third observation period during which customer churn is observed as demonstrated in Figure 1 (Appendix A).

To avoid left-censored data and properly capture customers' usage patterns, the dataset only includes customers that were active at the beginning of the study and omits any new entry into the study following the origin time. In addition, to capture and compare customers' usage behavior and variations, the sample includes subjects that have at least 1 follow-up during the second period and omits customers for which the churn event occurred during the first usage observation period. Therefore, the final dataset includes 11 647 subjects, of which 759 subjects have canceled the pay-TV subscription during the 6-month churn observation period. Lastly, all statistical analyses were conducted using SAS® Enterprise Guide 7.1 statistical software. Precisely, the first step of the study relied on the LIFETEST procedure in SAS as it is responsible for estimating and plotting Kaplan-Meier survival curves. The second step of the study and model building relied on the PHREG procedure statement in SAS as it generates the semi-parametric analyses using the Cox Proportional Hazard model.

VAS user-group definition and clusters

Users were clustered according to their VAS usage frequency and behavior between both usage observation periods (i.e. usage period 1 and usage period 2). Firstly, overall free VAS and payable VAS usage for each subject was summed for each period and then ranked into deciles. Similarly to linear regression, survival models are very sensitive to outliers and ranking usage into deciles decreases the influence outliers may have on the survival and hazard functions. For each usage period, subjects without recorded usage were assigned a rank of "0" (i.e. non-user) while subjects with recorded usage were ranked into deciles (1 to 10). Afterwards, subjects were clustered according to their sequential usage ranks between both usage periods to create user groups that capture usage frequency (light, medium, heavy users) as well as usage patterns between the two periods (i.e. increasing, constant, decreasing). In addition, as suggested by Prinzie and Van den Poel (2006), clustering subjects according to their sequential usage pattern will also increase the models predictive accuracy.

A two-step approach was used to cluster subjects into user-groups representing the different VAS user types and patterns. The first step was to explore the number of clusters necessary to accurately group users while maintaining sufficient variability in usage behaviors. Using the PROC cluster statement in SAS, Ward's clustering method was used to identify the most optimal number of clusters as it is recognized to produce very compact clusters and minimize within-cluster variance, thereby maximizing the difference between clusters. The second step was to assign subjects to different clusters by using the optimal number of clusters (i.e. seeds) from Ward's method as an input into the K-means clustering.

RESULTS

Clusters and VAS user-groups

The results summarized in Table 4 and Figure 2 (Appendix B) indicate a few clustering alternatives. The first alternative is to select 6 clusters, as the first relatively large peak in the pseudo T-squared statistic plot (Figure 2 – Appendix B) is at its highest point at 5 clusters, and according to the selection criteria, moving back 1 cluster from this peak indicates a good clustering option. In addition, the Cubic Clustering Criterion (CCC) value of 5.29 is above the minimum threshold for well-defined clusters (i.e. CCC > 2.0) and this alternative also accounts for 91.1% of the variation (R^2=91.1). The second option is to select 11 clusters, as the pseudo T-squared statistic plot (Figure 2 – Appendix B) depicts another notable peak at 10 clusters, and moving back from this peak indicates another good clustering option. The CCC is 6.60 and this clustering option accounts for 95.2% of the variation (R^2=95.2). However, even if both of these clustering options are good alternatives from a theoretical standpoint, these choices are not as sound for the purpose of this study or from a managerial standpoint. Firstly, the 6-cluster option does appear as a suitable and parsimonious choice, however this option does not generate a sufficient number of user-groups to account for individual's usage frequency (i.e. non-users, light-users, medium-users and heavy-users) and usage patterns (i.e. increasing, decreasing and constant). Consequently, this clustering option will limit the study's ability to fully test the hypotheses regarding VAS usage behavior. Secondly, although the 11 clusters generate more user-groups and variability to test the hypotheses, the relatively high number of user-groups will affect managerial contributions by limiting the study's interpretability and applicability in a business context. That is, even if the 11-cluster option can be argued as the best clustering option from a theoretical perspective, it is not exactly parsimonious and makes interpretation significantly more difficult and complex from a managerial perspective. Although the pseudo T-squared statistic plot does not depict another peak between 5 and 10 clusters, the next significant increase in the pseudo T-squared statistic recognized in Table 4 (Appendix B) is at 8 clusters. As suggested by the selection criteria, moving back 1 cluster to 9 clusters is the best alternative to balance theoretical selection metrics, provide more parsimony and interpretation for this study and its managerial implications. In fact, this clustering option yields a CCC value of 7.54 and is well above the minimum threshold required for a good clustering option and also accounts for 94.2% of the variation ($R^2 = 94.2$). In addition, even if this clustering option does not follow all of the theoretical recommendations regarding cluster selection, this option only reduces the ideal amount of clusters by 2 and is not expected to generate significantly different findings than would the 11-cluster outcome. Therefore, for the purpose and objectives of this study, the 9-cluster option provides the necessary level of granularity to test the hypotheses parsimoniously without limiting the study's managerial implications.

Following the selection criteria on the number of user-groups, clusters were then qualified and interpreted according to their cluster means for both usage observation periods summarized in Table 5 (Appendix B). Because subjects were clustered by their sequential average usage ranks, the cluster means represents subjects' usage ranks for each usage period. Therefore, by comparing cluster means for both periods, we can qualify whether subjects within each cluster have, on average, increased, decreased or held usage constant during the observation periods. For each VAS usage observation period, clusters with means < 1st decile are considered non-users, clusters with means ≥1st and ≤ 3rd decile are light-users, clusters with means ≥ 4th and ≤ 7th decile are medium-users, and clusters with means ≥8th decile are heavy-users.

Kaplan-Meier Survival Curves

The initial results of the study show each of the VAS user-group's survival functions and corresponding Kaplan-Meir (KM) survival curves. The survival curves depicted in Figure 3 (Appendix C) represent each of the user-group's ordered failure times beginning with a survival probability of 1, then diminishing as the analysis moves from one ordered failure time to the next (Kleinbaum & Klein, 2012). According to the results in Figure 3 (Appendix C), there is distinctive evidence that the different user-groups or clusters exhibit different survival patterns.

Although the differences between some of the survival curves are minimal, the Log-Rank test indicates whether or not the KM curves displayed in Figure 3 (Appendix C) are statistically different from one-another. Given the Log-Rank p-value <0.0001 (Appendix C – Table 9), the null-hypothesis suggesting that there is no overall difference between survival curves is rejected and we can conclude with very high confidence that the KM curves for the user-groups are statistically different from one another. According to these results, the survival plot for Cluster 5 is consistently higher than any of the other user-groups. This indicates that subjects in Cluster 5 (i.e. medium-users with an increasing trend) have a better survival prognosis than any other user-group in the study, including the control group of non-VAS users (Cluster 1). In fact, according to the censorship summary presented in Table 10 (Appendix C), 97.6% of the observations in Cluster 5 are censored which means that the churn event during the study period only occurred for 2.4% of the subjects. The second user-group with the lowest number of failures during the study period (censorship = 96.10%) is Cluster 9 (i.e. medium-users with a constant usage trend). The third user-group with the lowest level of churn and greater survival probability is Cluster 2 (light-users with an increasing trend). In comparison, Cluster 1 (i.e. non-users of VAS) has a lower survival distribution than Clusters 5, 9, and 2 suggesting that users with constant and increasing usage of VAS are indeed less likely to churn than non-users. On the other hand, Clusters 4, 6, and 8 have a lower survival prognostic than non-users and Clusters 3 and 7 appear to have similar survival functions than non-users. Although some of these results provide insights into different churn patterns among VAS users and non-users, variations in some of the survival functions are relatively small. In addition, the Kaplan-Meier estimator only predicts subjects' survival distribution based on how failures are distributed over time and does not consider the effects explanatory variables have on the survival function or on a customer's likelihood to churn (i.e. hazard function). Nonetheless, there is evidence to suggest that VAS usage does have negative effects on churn and support some of the study's hypotheses. For instance, these initial results show that several VAS user-groups are less likely to churn than non-users (H2a), users with increasing or constant VAS usage frequency churn less (H3a), with the exception to the heaviest of users (H3c). However, these initial results do not account for the influence other predictors have on subjects' survival function and hazard function. Therefore, results from the Cox Proportional Hazard model will be used to better depict customers' churn likelihood and further support these initial findings.

Cox Proportional Hazard Model

Results generated by Cox Proportional Hazards show each of the covariates' beta coefficients, hazard ratios (HR), as well as their significance level. Although the Cox Proportional Hazard model does yield beta-coefficients for each predictor variable, the measure of effects is typically done with the hazard ratio because it displays the true risk that the event will occur for each unit of time. That is, the hazard ratio will describe the relationship between a given covariate and a subject's survival time while the beta coefficient describes the relationship between the covariate and the hazard rate. However, before interpreting the covariates' effects, the model needs to be assessed for overall goodness of fit and proportional hazards. Generating the best fitting model requires goodness of fit tests, outlier detection, linearity and functional form validation and lastly, proportional hazards tests (Wilson, 2013).

Baseline Cox model and fit diagnostics

Results for the baseline Cox survival model are summarized in Table 11 (Appendix D). The Akaike Information Criterion (AIC) is a common approach for comparing the overall fit of survival models (Kleinbaum & Klein, 2012) and as demonstrated in Table 12 (Appendix D), the baseline model has an AIC fit statistic of 157 706.49. To further explore and potentially improve model fit, the baseline model is tested for influential outliers and whether or not the functional form is indeed linear.

Firstly, deviance residuals are plotted to test for the presence of outliers and whether or not the data adequately fits the model. When a model has good fit, deviance residuals are symmetrically distributed around 0 and relatively large residuals indicate that certain observations have bad fit and are potentially considered outliers (Fitrianto & Jin 2013). When censoring is minimal (<25%), deviance residuals will be normally distributed while datasets with censoring statistics greater than 40%, a large mass of residuals near 0, but the normal distribution will be distorted (Therneau et al., 1990). Deviance residuals plots are summarized in Figure 4 (Appendix D) and given that the dataset has over 90% censorship, the normal distribution is distorted as expected. However, the deviance residuals do appear to follow somewhat of a symmetrical pattern, but do nonetheless appear to lie far from 0 indicating potential problems with influential points. To test whether or not the data distribution has overly influential data points, DFBETA are plotted for all continuous covariates. Generating DFBETA's in survival analysis will produce a plot that displays the estimated change in regression coefficients upon deleting each observation for the given predictor variable (Fox, 2011). The cut-off value for influential observations is 1 or $2/\sqrt{n}$ where n is the number of observations in the dataset (Belslery et al. 1980). The dataset includes a total of 137 786 observations for the 11 647 subjects and thus, the cut-off value for this study is 0.005. The DFBETA plots are summarized in Figure 5 (Appendix D). Despite the fact that deviance residuals are not tightly distributed around 0, the DFBETA plots for most variables do not lie beyond the 0.005 cut-off value. However, AGE and SUB_DURATION do appear to have overly influential data points as several observations lie just above the acceptable cut-off value but are still below the normally accepted cut-off value of 1. Although the influence of outliers is minimal, findings for the deviance residuals and DFBETA values for AGE and SUB_DURATION may require treatment to better fit the model.

Furthermore, similarly to linear regression, the proportional hazard model assumes that relationships between covariates and the hazard ratio are linear, and non-linear relationships mean that the interpretation of the hazard ratio will be incorrect (Wilson, 2013). A common approach to testing for linear relationships in survival analysis is to examine the functional form of covariates included in the model. Testing functional form is done by plotting martingale residuals against each of the covariates and fitting a loess regression line to determine whether the relationship looks linear, quadratic, or threshold (Wilson 2013). Thus, martingale residual plots and loess lines are generated for all continuous variables in the model and are summarized in Figure 6 (Appendix D). With the exception of OVERDUE_BILL, the loess lines indicate that relationships for the covariates in the model are indeed linear. However, this linearity assumption is violated for OVERDUE_BILL.

Although most DFBETA and functional form diagnostics fall within the acceptable parameters, patterns in the deviance residuals plots, the overly influential data point for SUB_DURATION and AGE, and lastly, the non-linear relationship for OVERDUE_BILL suggest that the model is not adequately fitted and requires transformations to manage these irregularities and improve model fit.

Transformed Cox Model and Proportional Hazards

To minimize the effects of outliers and improve functional form for OVERDUE_BILL, the natural log was derived for each continuous variable originally included in the baseline model. The martingale residuals plots and corresponding loess lines in Figure 6 (Appendix E) shows that functional form for OVERDUE_BILL becomes linear and the remaining continuous variables maintained their linearity despite the log-transformation. The log-transformed model also shows a significant drop in AIC from the baseline model suggesting that the corrective measures significantly improved model fit. In fact, the AIC fit statistic dropped from 157 706.49 in the baseline model to 125 231.03 in the log-transformed model (Appendix E – Table 14). In addition, deviance residuals summarized in Figure 8 (Appendix E) are also much more symmetrical suggesting better fit to the transformed model.

The last step in the model's fit diagnostic is to determine whether or not the proportional hazards assumption holds. The proportional hazard assumption is likely violated for time-dependent variables, but it must hold for CLUSTER, which is in fact time-independent. To test proportional hazards, correlations between Schoenfeld residuals for each covariate and the ranking of individual failure times are tested (Kleinbaum & Klein, 2012). The proportional hazard assumption for a given covariate is respected when correlations are near 0 (Kleinbaum & Klein, 2012). In addition, the null hypothesis (i.e. proportional hazards are not present) is rejected for correlations with significant p-value at 95% confidence suggesting that there is sufficient evidence to conclude that proportional hazards are indeed present for the given variable. As expected, the proportional hazard assumption is violated for several of the time-dependent predictors (LIVE_TV_USG, ALL_FVAS_USG, STB_OWN, REGION, HOME_TYP) as each of these variables have relatively high Pearson correlation coefficients and significant p-values (Appendix E – Table 15). However, most importantly, the results do indicate that the proportional hazard assumption is respected for CLUSTER, the only time-independent variable. Because the dataset includes a time-varying predictor for each monthly period, the PROC PHREG procedure in SAS does account for the variation in predictor variables (SAS® Institute Inc., 2016). However, when the proportional hazard assumption does not hold, hazard ratios will not account for the influence the predictor variable has at a specific time, but will estimate an average hazard ratio for the study period (Schemper, 1992). This may lead to inaccurate averages because violations at the beginning of the study may cause the average hazard ratio to be overestimated, while risks at the end of the study may be underestimated (Schemper, 1992). In addition, even if violations in the proportional hazard assumption is accepted for time-dependent variables, ignoring non-proportional hazards can lead to incorrect results and reduce the model's fit to the data (Ata & Söker, 2007). In fact, in the presence time-dependent data and non-proportional hazards, Ata and Söker (2007) found that interaction models (i.e. Extended Cox Model) perform better and provide better fit than models that ignore the assumption violation. Therefore, to better fit and estimate the effects of time-dependent variables violating the assumption (Ata & Söker, 2007), the Extended Cox Model is used to include the original time-dependent covariates and a product of these covariates with a function of time (i.e. LIVE_TV_USG, ALL_FVAS_USG, STB_OWN, REGION, HOME_TYP).

Extended Cox Model

Several different functions of time can be used to generate time interactions in the Extended Cox Model. Allison (1995) suggests that the PHREG procedure in SAS is sufficiently robust and a simple linear function of time should be chosen (Borucka & Poland, 2013). Therefore, in addition to the covariates included in the transformed Cox Model, 5 other variables were added the model to account for the interaction between the linear function of time (i.e. duration) and the covariates with non-proportional hazards (i.e. itLIVE_TV_USG, itALL_FVAS_USG, itSTB_OWN, itREGION and itHOME_TYP). In addition, to keep the most influential covariates in the extended model, backward elimination is applied to the procedure to automatically eliminate insignificant variables from the final model (p-value> 0.05).

Results for the Extended Cox Model indicate that the interaction terms did indeed improve model fit as the AIC statistic has reduced to 124 817.38 (Appendix F – Table 16). The interaction terms itLIVE_TV_USG and itALL_FVAS_USG are each significant which suggest that non-proportional hazards are indeed present and accounted for in the extended model. However, the interaction terms for itSTB_OWN, itREGION and itHOME_TYP are insignificant at a 0.05 level and were eliminated from the model. Therefore, after correcting to proportional hazards for itLIVE_TV_USG and itALL_FVAS_USG, the proportional hazards assumption now holds for itSTB_OWN, itREGION and itHOME_TYP. In fact, testing for significant covariates that interact with time is another method to test whether or not variables have different effects overtime and thus violate the proportional hazard assumption (Kleinbaum & Klein, 2012). Therefore, after correcting for outliers, functional form, and accounting for non-proportional hazards, the Extended Cox Model is the final model used to measure and interpret the effects covariates have on customers' hazard and survival time.

Extended Cox Model Interpretation

The results in Table 19 (Appendix F) include the model's beta coefficients, hazard ratios and significance levels. The hazard ratio depicts the relationship between the different covariates and the hazard function. That is, while other variables are held constant, the hazard ratio for a specific covariate displays the risk between that variable and the rate or potential the event will occur. The null value for this "exposure-outcome" relationship (i.e. hazard ratio) is 1, meaning that the covariate has no effect, while a hazard ratio greater than 1 suggests that hazard risk is increasing and a hazard ratio less than 1 suggests that the hazard risk is decreasing with each unit increase in the predictor variable (Kleinbaum & Klein, 2012). That is, for continuous predictors, the hazard ratio for a specific variable represents the percentage increase or decrease in the risk associated to a unit change in that variable. For categorical variables with several levels or classes, the hazard ratio is interpreted as the risk potential the group has in comparison to the reference group.

Firstly, the model's significance summary presented in Table 17 (Appendix F) shows that most of original covariates included in the baseline model are now significant at 95%. Although some of the covariates were not significant in the baseline model, corrections improved model fit and its ability to capture the different variables' effect on the hazard function. However, the results in the final model (Appendix F - Table 19) also show that effects for some of the clusters (i.e. user-groups) are insignificant. The first notable result is that LIVE_TV_USAGE (i.e. basic service usage) has a significant negative influence on the churn event ($\beta = -0.35320$, $p < 0.0001$) with a hazard rate of 0.702. This result means that for every increase in basic service usage, a customer's hazard rate will decrease by 29.8%. Therefore, there is sufficient evidence to support H1 and the hypothesized negative relationship between basic service usage and churn likelihood. Furthermore, the results also summarize the level of risk cluster 2 to 9 display in reference to the control group of non-users (i.e. cluster 1). Although CLUSTER as an overall predictor is significant at a level of 0.0001 (Appendix F – Table 18), the risk potential for some of the clusters (i.e. user-groups) are insignificant. In fact, several of the insignificant clusters are also those that appeared to have similar survival characteristics in the Kaplan Maier survival plot in Figure 3 (Appendix C). Nonetheless, clusters 2, 5 and 9 display significantly lower hazard ratios when referenced to the control group of non-VAS users (i.e. cluster 1). In fact, light-users with an increasing VAS usage frequency (cluster 2) have a hazard ratio of 0.683 ($p<0.0001$), which means that this cluster exhibits 31.7% less risk control group. Medium users with a constant VAS usage trend (cluster 9) exhibit 24.9% less risk than the control group (HR=0.751; $p<0.0001$) and medium users with an increasing VAS usage trend (CLUSTER 5) show 57.8% less risk than the control group (HR=0.422; $p<0.0001$). In addition, even if overall VAS usage has been accounted for when defining the clusters, the results show that increases in free VAS usage frequency reduces churn risk while payable VAS has nearly no effect to a slight positive effect on churn risk. That is, for every unit increase in FVAS (ALL_FVAS_USG), a user's hazard decreases by 8.3% (HR=0.917; $p<0.0082$) while PVAS usage (ALL_PVAS_USG) has a very minor but positive effect on a customer's hazard risk (HR=1.060, $p < 0.0001$). Therefore, in addition to Kaplan Maier survival curves presented in Figure 3 (Appendix C), the hazard ratios for clusters 2, 5, and 9, the effects of free-VAS usage, there is sufficient evidence to support H2a and deduce that consistent VAS users are less likely to churn than non-users. In addition, the hazard ratio for ALL_PVAS_USG is very close to the null value, suggesting that it has little effect on churn risk. Contrary to the effect of FVAS usage, these results suggest that for every increase in PVAS usage, a customer's risk increases by 6%. Therefore, given that FVAS does indeed have a greater influence on reducing churn risk than PVAS, H2b is also supported. However, because the effects of some the VAS user-groups are insignificant, H3a can only be partially supported. The results do show with a very high confidence that user-groups with increasing or consistent usage frequency are less likely to churn. Precisely, medium users with an increasing usage pattern (CLUSTER 5) are less at risk than constant medium VAS users (CLUSTER 9) and even light users (CLUSTER 2). However, H3a cannot be fully supported because two of the user-groups show insignificant effects despite their increased or constant usage patterns (CLUSTER 4). In addition, H3b cannot be supported because the effects for user-groups with decreasing usage patterns (CLUSTER 3 and CLUSTER 7) are insignificant. Also, the effects of the heaviest VAS users (CLUSTER 6) are not influential in the model. However, the results depicted by the Kaplan Maier curves in Figure 3 (Appendix C) as well as the proportion of censored observations summarized in Table 10 (Appendix C) do clearly demonstrate that among any of the VAS user-groups, heaviest users are those that have the lowest survival prognostic and the highest level of churn during the study period. Therefore, even if the effects of cluster 6 are insignificant in the final model, the survival curves depicted by the Kaplan Maier estimator are indeed significant and provide sufficient evidence to support H3c and deduce that heaviest users do indeed experience the greatest churn risk.

Furthermore, all of the control variables included in the model generated significant results. Firstly, for subscription variables, total revenue (TOT_REV) has a positive influence on a customer's hazard rate ($\beta= 0.17341$, $p < 0.0001$) and for every unit increase in customer spending, a customer's hazard rate increases by 1.89% (HR=1.189). In addition, overdue bill amount (HR=1.141; $p<0.0001$), TV discount (HR=1.211; $p<0.0001$), and the customer's TV subscription duration (HR=1.250; $p<0.001$) are also recognized to contribute to a customer hazard risk and churn potential. That is, increases in overdue billing amounts, discounts and the TV subscription duration will each increase a customer's churn potential. On the other hand, without regard to the type of services a customer has subscribed to, the longer a customer's relationship with the operator (SUB_DURATION), the least likely they are to experience churn (HR=0.568; p<0.0001). In addition, the number of services a customer has subscribed to (QT_SERVICES) will also gradually decrease the risk to churn. In comparison to the reference group that are only subscribed to the TV service (QT_SERVICES=1), every additional service subscription will decrease a customer's hazard ratio. For example, when referenced to single TV subscribers, customers with two services (QT_SERVICES=2) have 26.9% les risk (HR=0.731; p<0.0001), customers with three services (QT_SERVICES=3) have 46.7% less risk (HR=0.533; p<0.0001) and customers with 4 services (QT_SERVICES=4) have 56.6% less risk (HR=0.434; p<0.0001). Also, in support to findings on VAS usage, customers that subscribe to premium services (PREMIUM_SERV) have 17.4% less risk compared to customers that are not subscribed to premium services (HR=0.826; p<0.0001). Lastly, the ownership status of the customer's set-top box also has an effect on churn likelihood. That is, customers that lease the set-top box (STB_OWN=0) rather than purchasing and owning it have a 24.8% greater risk to churn than the reference group that own the hardware (HR=1.248; p<0.0001). However, customers that have multiple set-top boxes of which at least 1 is owned (STB_OWN=2) have 23.4% less risk than the referenced group that have a single owned set-top box (HR=0.766; p<0.0001).

Lastly, all demographic variables included in the model also have significant effects on the model's hazard ratios. Firstly, age has a significant influence on the hazard rate (HR=0.432, p-value<0.0001) and every unit increase in age decreases the hazard rate by 56.8%. In addition, customers in a rural region (REGION_0) have 17.2% less risk than the referenced group living in urban regions (HR 0.828, p-value<0.0001). Lastly, customers in multi-dwelling homes (HOME_TYPE_0) are 15.7% more likely to churn than referenced customers in a single-family home (HR 1.157, p-value<0.0001).

DISCUSSION

The research set out to understand whether or not VAS offered by pay-TV providers truly reduces customers' likelihood to churn. The first objective of the study was to determine if VAS users were indeed less likely churn than non-users and secondly, identify how different VAS usage patterns influenced churn likelihood. Although the current body of churn literature for telecommunications services has identified various predictors that influence customer churn, very few studies have considered and successfully measured the different effects value-added services have on churn likelihood. Even if usage has been recognized as one of the most important predictors of churn (Anh et al., 2006), financial measures for usage is not applicable to fixed cost subscription services and does not distinguish the effects between basic service usage and value-added service usage. Thus, the main contribution of the study is to provide a much broader perspective on how service usage, especially VAS usage and various usage behaviors contribute to reducing customer churn. Although certain relationships of interest in the final research model were insignificant, the overall results of this study provide significant insight to answering each of the study's research questions.

Beginning with the first research question (i.e. *are users of value-added services less likely to churn than non-users?*), results from the initial Kaplan-Meier estimator survival plots and the final hazard model provide evidence that most consistent VAS users are indeed less likely to churn than the control group of non-users of VAS. As hypothesized, several of the VAS user-groups exhibited a greater survival prognostic than the control group of non-users. In fact, light but increasing users of VAS (cluster 2), medium VAS users with an increasing usage trend (cluster 5) and medium users with consistent usage (cluster 9) exhibited significantly lower churn risk than the control-group of non-VAS users. In addition, the Kaplan-Meier survival plots show that these same user groups have the greatest survival prognostic than any other of the user-groups. These results do show that VAS users with different usage patterns also have different churn risks and behaviors. However, results for light-users that have significantly increased their usage to become medium users (cluster 4) are actually more susceptible to churn than the control-group and nearly as likely to churn that the heaviest of users (cluster 6). However, contrary to the significant findings for the other consistent or increasing users of VAS, cluster 4 only includes 188 of the 11 647 subjects; the least amount of subjects compared to any of the other user-groups (Appendix B – Table 6). In addition, the effects for cluster 4 were found to be insignificant in the final Cox Proportional Hazard model. Therefore, even if the findings for cluster 4 are not as expected or hypothesized, the findings for the three other user-groups with increasing and consistent VAS usage show that VAS usage does in fact have negative influence on churn likelihood. These results confirm findings that Madden et al. (2006) could not establish and provide actual behavioral insight to confirm the influence VAS has on customers' intentions to maintain its relationship with a service provider (Santouridis & Trivellas, 2010). Contrary to Madden et al. (2009) and Santouridis and Trivellas (2010), the study did not rely on cross-sectional data and customer intentions. Rather, the research relied on actual customer behavior and customer churn outcomes to estimate churn risk using survival analysis. These findings also show and support that behavioral variables provide better statistical predictability (Li et al. 2015), especially when service usage is not directly correlated to customer expenditure.

Furthermore, the findings also show that the type of VAS is also important when examining the influence these services have on customer churn. In fact, excessive use of payable VAS in mobile service has been recognized to increase churn risk (Geetha & Kumari, 2011). Even if the study did not explicitly examine frequency of use for payable VAS, the research does show that VAS usage frequency from services that generate additional customer expenditure increases churn risk. Given that free VAS usage statistics exceeds payable VAS within the sample population (Appendix 3 – Table 3), the slightly yet positive relationship on churn is likely offset by the much stronger effects of free VAS usage. In fact, given that free VAS can be considered as a free premium service associated to the basic pay-TV subscription fee, these findings validate that free premium services do indeed have a positive effect on customer loyalty (Unhanandana & Wattanasupachoke, 2012). Also, in support to Zeithaml's (1988) value and tradeoff concept as well as Jamal and Bucklin's (2006) notion of payment equity, it is reasonable to suppose that free VAS generates more value for customers because the tradeoff between the benefits and sacrifices is more important for payable VAS than free VAS. Although free VAS was expected to have a more important beneficial relationship on churn than payable VAS, the positive influence payable VAS has on churn is unexpected. Even the effect of payable VAS is relatively small, this result corroborates with the fact that customer expenditure is positively related to churn (Anh et al. 2006; Burez & Poel, 2007; Portela et al. 2010). These relationships suggest that customers that generate payable VAS usage through transactional purchases or premium subscriptions services may be more price-sensitive and more likely to seek alternatives. However, contrary to Jamal and Bucklin (2006), it would appear that the benefits or additional utility derived from payable VAS does not compensate for the effects higher expenditure has on churn. However, the results also show that customers that have subscribed to premium services (i.e. premium channels and subscription video-on-demand services) exhibit less churn risk than those that have not subscribed to premium services. That is, the results suggest that the influence of transactional VAS and subscription-based VAS may not be the same. One possible explanation for this relationship is that subscription-based VAS (i.e. premium channels, subscription video-on-demand) has a different relationship on churn than transactional purchases of VAS (i.e. transactional video-on-demand). That is, it is possible that expenditure related to the monthly subscription cost falls within customers' payment expectations while transactional purchases go beyond the recurrent subscription fee. In relation to Jamal and Bucklin's (2006) notion of payment equity, it is also possible that premium subscription services generate more value than transactional purchases, thereby compensating for the negative effects related to the increase in premium subscription fees. However, because usage from payable subscription services and transactional services are not explicitly measured, these different effects cannot be interpreted beyond the current results. Conversely, findings from Geetha and Kumari (2011) may provide another explanation such that excessive users of payable VAS may become more susceptible to churn when payable VAS exceeds a certain threshold and proportion of the monthly service cost. That is, heavier usage of payable and transactional VAS may be related to more churn events than light to moderate usage of payable VAS. Nonetheless, even if the study did not isolate all different types of payable VAS, these results do show that there is an important distinction between effects of free and payable VAS and potentially the type of payable VAS as well.

Furthermore, the study shows that the availability of VAS is not only important when subscribing to a new service, (Krishnan & Kothari, 2006; Ko et al., 2013; Ku et al., 2009) but also an important component that contributes to reducing customer-switching behaviors. However, the results also show that the influence VAS has on customer churn is not a simple binary relationship and VAS users are not systematically less likely to churn than non-users. That is, because VAS user-groups with different underlying usage patterns exhibit different survival prognostics and risks, the results provide compelling insight to help answer the second research questions (i.e. *how do different VAS usage behaviors influence a customer's likelihood* *to churn?*). Although some of the effects for user-groups and clusters in the final Cox Proportional Hazard model were insignificant, the Kaplan-Meier survival curves combined to the significant user-groups in the final model demonstrate that users with different usage patterns do in indeed exhibit different survival patterns.

For instance, the final Cox model shows that users with an increasing VAS usage frequency (i.e. cluster 2 and cluster 5) exhibit less risk to churn than consistent users (cluster 9). Also, the Kaplan-Meier survival curves show that clusters with a decreasing usage frequency have similar survival prognostic than the control group (cluster 3 and cluster 5) while others (clusters 6, 7 and 8) were even inferior to the control group. Although some of the hypotheses regarding usage frequency could not be fully supported by the final Cox PH model, results between the Kaplan-Meier estimator as well as the final Cox model do provide evidence that different VAS usage behaviors are associated to different levels of risk. In support to Keaveney and Parthasarathy (2001), the results do show that increased usage frequency does have a positive effect on service continuation. That is, medium users with an increasing usage trend were the least likely to churn compared to light-users. In fact, the results show that moderate VAS usage has the most beneficial effect on customer churn behavior (i.e. medium users). This finding is also consistent with the argument that customers with heavier VAS usage frequency have greater attachments to the service (Lee et al., 2001). However, because the heaviest of VAS users, exhibit the greatest proportion of churn events and were the least likely to survive, there may be a certain threshold where increased usage frequency actually stimulates switching behaviors. Although VAS user groups were defined using overall VAS usage, this finding agrees with Geetha and Kumari's (2011) findings that excessive VAS spending was positively related to churn. However, even if this finding supports Geetha and Kumari (2011), the explanation behind this behavior may be different. This difference is explained by the fact that Geetha and Kumari (2011) explicitly examined VAS spending compared to basic service spending while this study defined VAS user-groups according to overall VAS, whether payable or free. Although the heaviest of VAS-users in this study might also be heavy users of payable VAS, Anh et al. (2006) suggest that heavy users become most familiar with the service, have greater expectations and actually become more likely to explore and try more advanced alternatives. Value-added services offered by pay-TV operators are designed to create more value, support new customer viewing behaviors and ultimately retain customers from switching to a competing more advanced service. Although the results show that the various value-added services do have a beneficial effect on customer retention, it is not unreasonable to assume that accumulated usage and experience may increase customer expectations and entice heavy users to explore and try other potentially more advanced alternatives. In fact, these findings also support Dover and Merthi's (2006) argument that customer experience and knowledge can be both beneficial and detrimental as increased awareness and knowledge can favor the adoption of competing services. This may be especially true for heavy users of online value-added services because such users may have more exposure, willingness and capabilities to try new online service alternatives. That is, customers that use pay-TV operators' online-TV applications may become more comfortable with mobile phones, tablets and other Internet connected devices making them more likely to try new online alternatives that compete with the pay-TV product. Another explanation in support to Geetha and Kumari (2011) is that heaviest users of VAS may also become price-sensitive because of increased customer spending whether from the basic TV service, premium subscription services, and transactions value-added services and therefore seek alternatives to lower costs.

In addition to usage frequency, the results also show that customers' usage pattern (i.e. constant, increasing or decreasing) is also an important component to understanding the relationship between usage and churn likelihood. Although the hypotheses could not be fully supported by the final Cox model, the Kaplan-Meier estimator and survival plots provide evidence that churn behavior is also related to different VAS usage patterns. In support to Anh et al. (2006), Allenby (1999) and Zorn et al. (2010), the results show that customers do not suddenly churn without behavioral evidence such as usage status (i.e. "super-active", "active" and "non-active"). That is, even if the customer is still an active user, decreasing VAS usage may be one of those behavioral indications that the customer is less committed and the relationship is at risk. According to the initial results and survival plots, subjects with decreasing usage patterns showed either a lower or similar survival probability than the control group. Light-users with decreasing VAS usage patterns (clusters 3 and 7) may in fact become non-users and share similar survival behaviors than the control group. That is, clusters 3 and 7 were initially light-users and further decreases in VAS usage might show that they rarely engage with VAS thereby eliminating the effects VAS has on customer relationships and loyalty. Medium and heavy-users of VAS that suddenly decrease usage have lower survival probabilities than the control group (i.e. clusters 6 and 8). The sudden decrease in VAS usage may suggest that these subjects have also decreased overall usage for both basic and value-added services. Supported by Anh et al. (2006), a possible explanation may be that these medium and heavier users have already begun adopting and trying new competing service alternatives and are ultimately becoming less engaged to the pay-TV service. That is, initial trials may cause these users to gradually transfer usage time to competing alternatives and reach a critical point where users move from a trial phase to a permanent adoption thereby canceling the pay-TV subscription service.

Lastly, in addition to findings regarding VAS, results for control variables included in the hazard model validate and even contribute to current recognized churn predictors for Pay-TV services and telecommunications services alike. For subscription variables, the model supports that customer expenditure (Anh et al. 2006; Burez & Poel, 2007; Portela et al. 2010) and overdue billing amounts (Anh et al., 2006; Jamal & Bucklin, 2006) do indeed have a positive influence on churn. However, contrary to Portela (2010) and in support to Anh et al. (2006), the results also show that service usage is in fact an important predictor of customer churn. In addition, contrary to (Geetha & Kumari (2011), the results show that customer expenditure is in fact an inadequate measure of usage, especially for pay-TV services. In fact, while customer expenditure does increase churn risk, the results also show that actual service usage frequency (i.e. in hours) has a negative and beneficial relationship to churn. Thus, in support to Kermati and Ardabili (2011) and Ascarza and Hardie (2013), the results show that basic service usage frequency has a negative and beneficial influence on churn. Because pay-TV service fees are not directly related to usage, as it may be the case for other telecommunications services, the results show that although customer spending is indeed an important churn predictor, researchers need to consider behavioral service usage metrics. In addition, the results support that service bundling (Jamal & Bucklin, 2006; Burez & Poel, 2007; Prince & Greenstein, 2014) and a customer's subscription length has a negative and beneficial effect on churn. However, the research distinguishes between the pay-TV subscription length and a customer's overall relationship length including other telecommunications service subscriptions. Surprisingly, the relationship between the two durations differ and the findings show that the subscription duration for customers that have only subscribed to the pay-TV service is actually positively related to churn likelihood. Results for

TV subscription length suggest that long-term customers may become sensitive to alternatives or promotional offerings from competing providers. However, the discrepancy between TV subscription duration and overall subscription duration may also be explained by an interaction with the quantity of services a customer has subscribed to with the operator. That is, because customers overall subscription duration accounts for other services, the effects of service bundling (Jamal & Bucklin, 2006; Burez & Poel, 2007; Prince & Greenstein, 2014) is likely responsible for this discrepancy. Lastly, the findings also show that a customer's set-top box ownership status is also an important and significant predictor of churn that should be included in future prediction models. In fact, the results show that customers that purchase and own the set-top box are less at risk than customers that lease the equipment. Because the set-top box is essential to accessing the digital-TV service, customers that have purchased and own the hardware are more attached to the service provider due to the initial cost and purchase of that hardware. When the set-top box is owned rather than temporarily leased, a customer that switches to a new service provider would lose this initial investment because the set-top box is typically proprietary to a single pay-TV service. Therefore, hardware ownership status behaves as an exit barrier for service operators.

The model also provides support for demographic variables associated to churn literature. In support to Prince and Greenstein (2014), age has a negative influence on customer churn which is in fact opposite to findings for mobile services. While younger customers may be less likely to switch due to mobile device switching costs, younger pay-TV subscribers are at greater risk and this risk diminishes with time. In addition, although Jamal and Bucklin (2006) could not establish a significant relationship, the results contribute to their study by demonstrating that both home type and a customer's region are indeed significant and important predictors of churn behavior. In fact, the results show that customers that live in multiple dwelling homes (i.e. apartment buildings) are at greater risk than those in single dwelling homes. In addition, customers living in urban regions are also more likely to churn than customers living in rural areas. Given that customers in apartment buildings are likely to move more frequently than more permanent single dwelling homes, they may be more inclined to shop and compare for alternatives more frequently and thus more likely to switch operators. For customers in urban regions, their risk level may be explained by the competitive dynamics and the availability of alternatives in urban regions. In addition, rural areas may have more limited access to broadband Internet and consequently, reducing access to new online competitors. Although these subscription and demographic predictors are not the focal point of the study, they do provide additional support and contribution to important predictors of churn for pay-TV services.

Theoretical implications

Firstly, the results of the study show that churn prediction models for telecommunications services should look at attributes beyond the core product in order to grasp a more thorough representation service usage has on customer churn. Although the research does support many of the currently recognized churn predictors, the results also show that research models need to look beyond subscription data, billing data and customer demographics, and leverage customers' actual behavioral characteristics. In an era of big data and digital services, churn prediction models are excluding a significant component to understanding customers' behavioral intentions and this the results of this study demonstrates that research models can benefit from examining customers' usage statistics and behaviors. Most importantly, the research demonstrates that these services are not only responsible to attract customers, but play an important role on maintaining existing customer relationships and customer loyalty. The findings provide a much broader perspective on the influences different types of value-added services have on customer churn, and most importantly, how different usage frequencies and patterns affect churn outcomes. Given that few studies have explicitly examined the influence VAS services and usage patterns have on customer relationships, the research sets the theoretical foundation on which future research can build and expand.

In addition, the results contribute to the literature dedicated to churn predictors relevant to pay-TV services and telecommunications services. The research has clarified contradicting effects usage has on customer churn, supported several recognized predictors, and most importantly, provided new behavioral predictors to improve churn prediction and better target high-risk customers. For instance, contrary to previous findings, the results show that customer expenditure is not an accurate measure for service usage and actual behavioral usage metrics are necessary to adequately evaluate the relationship between usage and churn. The research does support findings regarding the effects of customer expenditure. However, the results show that usage frequency reduces customer risk while increases in customer spending actually increase risk. Given that the research was interested in customer usage patterns for value-added services, the results also open a new direction of research to better understand how basic or overall service usage patterns may influence a customer's likelihood to churn.

Managerial implication

From a managerial perspective, the results show that pay-TV operators' efforts to extend the basic core service do have beneficial effects on maintaining their subscriber base. Given that customers are gradually shifting viewing time to other delivery methods (Nielson 2016), the results provide even more evidence to show that pay-TV operators need to capture this shift in content viewing and maximize customer engagement and time spent with their services. Firstly, while the results do indeed show that their efforts to extend the basic service does contribute to customer loyalty, they also show that these efforts should be focused on services that generate the most value to customers, even if it does not generate immediate revenue. While payable VAS does generate new income, operators will need to establish which provide greater returns; customer retention or additional ARPU. Although additional ARPU have been maintaining annual revenues for operators, this additional ARPU is likely to accelerate churn and further entice customers to seek alternatives.

Also, pay-TV operators need to closely examine customers' usage behaviors and patterns as they provide valuable insight to a customer's likelihood to churn. The results show that pay-TV operators can easily reduce customer churn by increasing the adoption and usage of its existing value-added service offering. However, the results also provide evidence that operators need to maintain and stimulate usage to maintain customer interest and engagement. Otherwise, customers that decrease usage may become more at risk as usage decreases. In addition, the results provide evidence that heaviest VAS users may be pre-disposed to switch and try new services. Although the research has not examined why heavy-users churn most, this result may be another motivation to continue to improve the service offering to maintain interest and usage for even the heaviest of users that may otherwise seek alternatives.

Lastly, the results also provide pay-TV operators with a much wider range of indicators and behavioral insight that can be relied upon to better target high-risk customers for proactive retention campaigns and improve overall retention performance.

LIMITATIONS AND FUTURE RESEARCH

Although the research provides significant contributions to academics and pay-TV operators, there are notable limitations to consider. Firstly, even if users were adequately segmented into user-groups with specific usage attributes, the highly censored nature of the data did limit variability and interpretation between some of the user-groups. In fact, according to a Canadian study (NGL Nordicity Group Ltd., 2016), monthly churn rates for telecommunications services in Canada vary between 1.3% and 1.6%. Thus, given that the research design recorded customer churn for a six-month period, average churn rates should vary between 7.8% and 9.6%; which is slightly higher than what was observed in the dataset. In fact, of the 11 649 subjects, 759 subjects churned during the study period which translates to a 6.5% churn rate. Therefore, 93.5% of the observations were censored thereby limiting the model's predictive ability. In fact, the deviance plots in Figure 8 (Appendix E) show that subjects with larger deviance residuals churned before the model could predict it (Gharbivand & Fernandez, 2008), meaning that predictive capability for churn events is limited. However, according to these same deviance residual plots, relatively small deviance residuals show that the model does adequately predict which subjects were less likely to churn and had longer survival times (Gharbivand & Fernandez, 2008). In addition, contrary to logistical regression, survival analysis does handle and interpret the effects of censored variables (Kleinbaum & Klein, 2012) and this explains why the model's performance for censored observations performed well. Nonetheless, highly censored data did limit the model's predictive accuracy for churned events. Even if there is a clear and significant distinction between the control group and some of the VAS user-groups, the final model was unable to yield significant results for some of the VAS user-groups, especially for user-groups with similar censorship distributions. In fact, the initial survival analysis did show that user-groups were significantly different from one-another, however, limited survival variability between some of the user-groups made it impossible for the final hazard model to capture some of these relationships. In addition, although the model did yield adequate clusters and user-groups, the number of subjects in some of the clusters may also be responsible for the insignificant findings. For instance, cluster 4 was assigned 188 subjects while subjects and their underlying usage patterns were more evenly distributed across other clusters. In addition, even if theoretical selection criteria suggested that 11 clusters was the best alternative, this limitation shows that this issue would not have been addressed with even more clusters. That is, an increased number of clusters may have generated even more insignificant user-groups with uneven subject distributions. The research could have further reduced the number of user-groups. However, this would have limited the study's ability to examine behavioral patterns. In addition, the fewer clusters would have likely generated less homogenous user-groups thereby limiting their interpretation. Therefore, even if the insignificant user-groups are indeed a limitation to the study, balancing theoretical and managerial criteria for cluster selection did nonetheless yield significant findings for several of the usage behaviors and patterns of interest and a larger number of user groups are not expected to alter the study's findings.

In addition, although the research was mostly interested on how overall VAS usage behaviors and patterns influenced customer churn, the study did not make a distinction between the types of VAS when examining these usage behaviors. Even if the research shows that free and payable VAS do not generate the same effects on churn, the study does not shed light on how the different usage behaviors related to different types of VAS influences customer attrition. In addition, the results even suggest that the type of payable VAS (transactional vs. premium service subscription) may have different influences on customer relationship duration. Therefore, to further understand the influence different types of VAS have on customer loyalty, future research should examine customer usage behaviors and patterns for each type of VAS (i.e. free vs. paid). These findings would provide valuable insight to better depict how different types of VAS influences the customer relationship.

Furthermore, because data was obtained from a single pay-TV operator, below average churn rates may be explained by above average performance and the operator's capability to retain customers. Another possibility may be that the 6-month churn observation excluded some of the more important seasonal fluctuations. In fact, this emphasizes another research limitation that concerns seasonal effects on both usage distribution, and churn distribution. Although the research design does capture usage patterns between the two most important broadcasting and programming calendars during the year (i.e. fall / winter), usage patterns may not be specifically attributed to customers' usage behavior, but to seasonal programming variations. Therefore, given the relatively low monthly churn rate and potential seasonal effects on usage distribution, future studies on usage patterns should span over a longer period of time to reduce censorship and better capture seasonal variations for both usage and churn distribution. This will improve the model's overall fit and its predictive ability to capture the relationships between VAS usage patterns and churn events. In addition, given that the dataset relied on a single pay-TV operator, replicating the study over a longer period of time with a different service provider will also help validate and even generalize the findings.

Secondly, because subscriber information included in the study is not profiled to individual users but rather to entire household, the research captures overall usage generated by all individuals in the household. Therefore, for customer accounts with multiple users, usage generated by VAS may not be representative of the account holder's usage behaviors. In fact, the research assumes that usage generated by individuals in the household will have an effect on the account holder's decision to sustain the TV subscription. In addition, a household may include multiple different socio-demographic profiles, which may have different media consumption behaviors. For instance, younger users may be more inclined to adopt new media delivery methods using Internet connected devices while older more traditional profiles may be more attached to the basic service delivery (i.e. live-TV). Therefore, for households with multiple users, the link between VAS usage in the household and the influence it has on the account holder is unknown. Moreover, even if users-groups were successfully clustered according to their specific usage attributes, the clusters do not take into account customers' socio-demographic profiles. Therefore, while usage patterns are significantly related to different churn patterns, there may be other influences at play that make certain user-groups more or less likely to churn. Considering other qualitative and socio-demographic profiling will likely generate even more homogenous user-groups. Therefore, future research should examine how individual user-profiles adopt and use value-added services and how this influences the decision making process to peruse, switch, or cancel the subscription service.

In addition, although the research did capture basic live-TV usage and several different types of value-added service usage, time-shifted TV using a digital recorder could not be captured. Although pay-TV providers offer this capability to its premium class of set-top boxes, playback for locally recorded content could not be measured. Because time-shifted TV goes beyond the basic service delivery, such usage would be measured as VAS. Given the variety of VAS usage that has been recorded for the study, results are not expected to vary by adding usage for recorded content. However, this does nonetheless impose a certain limitation to obtaining a complete and thorough picture of a customer's usage behavior and the complete portfolio of VAS available to pay-TV customers. In fact, even if the research was only interested in usage generated by the pay-TV service, future research could consider the effects competing service usage has on a customer's likelihood to churn. By considering a customer's overall time spent with media outside of the pay-TV offering, researchers can provide more insight on how customers' usage distributions vary across services and how this influences their engagement or likelihood to permanently adopt competing services and consequently cancel the pay-TV service. For example, subscribers that are also subscribed to or use content delivery services outside of the pay-TV subscription (i.e. Netflix, Youtube, AppleTV, etc) are exposed to competing offerings and perhaps more prone to cancel the traditional pay-TV subscription. While the research attempted to show how different usage patterns (i.e. constant, increasing, decreasing) with the pay-TV service influenced churn, understanding how overall content usage distributions vary across services can also be a very important indicator to predict when customers are most at risk. This would also enable pay-TV operators to identify a certain threshold whereby customers reduce their usage frequency and reach a tipping point at which they cancel their subscription. In addition, this can provide insight on which service attributes outside the pay-TV service generate value and engagement for customers. This research extension can also be valuable to explain how and why heaviest users of VAS become more likely to churn than any other user group.

In summary, to further improve managerial contributions, a more thorough understanding on customers' individual usage patterns across various different content and media distribution channels will better prepare traditional pay-TV operators to understand why customers are gradually abandoning traditional media delivery methods. Although subscriber losses are currently offset by increased ARPU (Advanced Television, 2016), a more thorough understanding on customers' service usage behaviors, whether from basic, value-added services and competing services will enable operators to better adapt the service offering, the business model and better protect the current subscriber base in a much more sustainable manner.

REFERENCES

Advanced-television (2016). Canada: Pay-TV subs decline, revenues steady. Retrieved from: http://advanced-television.com/2016/07/13/canada-pay-tv-subs-decline-revenues-steady/

Allenby, G. M., Leone, R. P., & Jen, L. (1999). A dynamic model of purchase timing with application to direct marketing. *Journal of the American Statistical Association, 94 (*446), 365-374.

Allison, P. D. (1995). Survival Analysis Using SAS®. a Practical Guide. Cary, North-Carolina: SAS Institute Inc.

Anh, J-H., Han, S-P., & Lee, Y-P. (2006). Customer churn analysis: Churn determinants and mediation effects of partial defection in the Korean mobile telecommunications service industry. *Telecommunications Policy, 30,* 552-268.

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

Ata, N., & Söker, T. (2007). Cox regression model with nonproportional hazards applied to lung cancer survival data. *Hacettepe Journal of Mathematics and Statistics, 36 (2),* 157-167.

Belsley, David A., Edwin Kuh, & Roy E. Welsch. (1980). Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley.

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., & Lemon, K., N. (1999). A dynamic model of customers' usage as anantecedent and consequence of satisfaction. *Jounral of Marketing Research, 36 (May),* 171-86.

Bolton, R.N., Lemon, K.N. & Verhoeff, 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,* 271-92.

Borucka, J., & Poland, W. (2013). Extension of the Cox Model for Non-Proportional Hazards Purpose.

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

Choi, C., Kim, C., Sung, N. & Park, Y. (2007). Evaluating the quality of service in mobile business based on fuzzy set theory. Fourth International Conference on Fuzzy Systems and Knowledge Discovery, Haiku, China, August 2007, 483-487.

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

Fox, J. (2011). Cox proportional-hazards regression for survival data. *An R and S-PLUS companion to applied regression*, 1-18.

Fitrianto, A., & Jin, R.L.T (2013). Several Types of Residuals in Cox Regression Model: An Empirical Study. *International Journal of Math, 7 (53),* 2645-2654.

Gharibvand, L., & Fernandez, G. (2008). Advanced Statistical and Graphical features of SAS® PHREG. In *SAS Global Forum 2008 Proceedings http://www2. sas. com/proceedings/forum2008/375-2008. pdf*.

Gardiner, J. C. (2010). Survival analysis: overview of parametric, nonparametric and semiparametric approaches and new developments. In *SAS Global Forum 2010. Statistics and Data Analysis*.

Gerpott, T.J., Rams, W., & Schindler, A. (2001). Customer retention, loyalty and satisfaction in the German mobile cellular telecommunications market. *Telecommunications Policy, 25 (4),* 249-69.

Geetha, M., & Kumari, J. A. (2011). Analysis of churn behavior of consumers in Indian telecom sector. *Journal of Indian Business Research, 4 (1),* 25-35.

Gronroos, C. (2004). The relationship marketing process: communication, interaction, dialogue, value. *The Journal of Business and Industrial Marketing, 19 (2),* 99-113.

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.

Jasrai, L. (2014). Predicting customer satisfaction towards mobile value-added services: an application of multiples regression. *The IUP Journal of Marketing Management, 13 (1).*

Kaplan E.L. & Meier, P. (1958). Nonparametric estimation for incomplete observations. *Journal of American Statistical Association, 53 (282).*

Kauffman, R. J., & Wang, B. (2008). Tuning into the digital channel: evaluating business model characteristics for Internet firm survival. *Information Technology Management, 9,* 215-232.

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. *Academy of Marketing Science, 29 (4),* 374-390.

Keramati, A., & Ardabili, S. M. S. (2011). Churn analysis for an Iranian mobile operator. *Telecommunications Policy, 35,* 344-356.

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

Kim, H-W., Chan, H. C., & Gupta, S. (2007). Value-based adoption of mobile Internet: an empirical investigation. *Decision Support Systems, 43,* 111-126.

Kleinbaum, D.G., & Klein, M. (2012). Survival Analysis: A Self-Learning Text (3rd ed.). New York : Springer.

Ko, H-T., Chang, C., & Chu N-S. (2013). An empirical investigation of the consumer demand for digital television application services. *Behaviour & Information Technology, 32 (4),* 397-409.

Ku, Y. F., Wu, C. M., & Deng, W. J. (2009). The relationship among service quality, perceived value, customer satisfaction, and post purchase intention in mobile value added services. *Computers in Human Behavior, 25 (4),* 887-896.

Kuo, Y. F., & Chen, P. C. (2006). Selection of Mobile Value-Added Services for System Operators Using Fuzzy Synthetic Evaluation. *Expert System with Applications, 30 (4),* 612-620.

Krishnan, R., & Kothari, M. (2008). Antecedents of customer relationships in the telecommunications sector: an empirical study. *Journal of Services Marketing, 6 (3),* 38-59.

Lai, F., Griffin, M. & Babin, B.J. (2009). How quality, value, image, and satisfaction create loyalty at a Chinese telecom. *Journal of Business Research, 62 (10),* 980–986.

Lai, Y., & Zeng, J. (2013). Analysis of customer churn behavior in digital libraries. *Electronic Library and Information Systems, 48 (4),* 370-382.

Lam, S.Y., Shankar, V., Erramilli, M.K., & Murthy, B. (2004). Customer value, satisfaction, loyalty, and switching costs: an illustration from a business-to-business service context. *Journal of the Academy of Marketing Science, 32 (3),* 293–311.

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.

Lim H., Widdows R., & Park F. (2006). M-Loyalty: Winning Strategies for Mobile Carriers. *Journal of Consumer Marketing, 23 (4),* 208-218.

Lin, T-C., Wu, S., Hsu J. S-C., & Chou Y-C. (2012). The integration of value-based adoption and expectation-confirmation models: an example of IPTV continuance intention. *Decision Support Systems, 54,* 63-75.

Liu, D. (2010). Proc Lifered or Proc Phreg (unpublished). Northwestern University, Chicago, IL.

Lu, J. (2002). Predicting customer churn in the telecommunications industry—An application of survival analysis modeling using SAS. *SAS User Group International (SUGI27) Online Proceedings*, 114-27.

Madden, G., Savage, S. J., & Coble-Neal, G. (1999). Subscriber churn in the Australian ISP market. *Information Economics and Policy, 11, 195-207.*

Nielson. (2016). The Nielson Total Audience Report, Q2 2016.
Retrieved from: http://www.nielsen.com/us/en/insights/reports/2016/the-nielsen-total-audience-report-q2-2016.html

NGL Nordicity Group Ltd. (2016). 2016 Price Comparison Study of Telecommunications Services in Canada and Select Foreign Jurisdictions.
Retrieved from : http://www.crtc.gc.ca/eng/publications/reports/compar/compar2016.pdf

Oliver, R. L., & Winer, R. S. (1987). A framework for the formation and structure of consumer expectations: review and propositions. *Journal of Economic Psychology, 8 (4),* 469-499.

Portela, S., & Menezes, R. (2010). An empirical investigation of the factors that influence customer churn in the Portuguese fixed telecommunications industry: A survival analysis application. *The Business Review, Cambridge, 14 (2),* 98-103.

Prince, J., & Greenstein, S. (2014). Does bundling reduce churn? *Journal of Economics & Management Strategy, 23 (4), 839-875.*

Prinzie, A., & Van den Poel, D. (2006). Incorporating sequential information into traditional classification models by using an element/position-sensitive SAM. *Decision Support Systems, 42(2)*, 508–526.

Ranaweera, C., & Prabhu, J. (2003). The influence of satisfaction, trust, and switching barriers on customer retention in a continuous purchasing setting. *International Journal of Service Industry Management, 14 (4),* 374-395.

Reinartz, W. J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing, 67 (1),* 77-99.

Sánchez-Fernández, R., & Iniesta-Bonillo, M. M. (2007). The concept of perceived value: a systematic review of the research. *Marketing Theory, 7,* 427-451.

Santouridis, I., & Trivellas, P. (2010). Investigating the impact of service quality and customer satisfaction on customer loyalty in Greece. *The TQM Journal, 22 (3),* 330-343.

SAS® Institute Inc. (2016). SAS Product Documentation (Modeling with Categorical Predictors)
Retrieved from:
https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_phreg_sect034.htm

SAS® Institute Inc. (2016). SAS Product Documentation (Time-Dependent Repeated Measurements of a Covariate)
Retrieved from:
https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_phreg_sect038.htm

Schemper, M. (1992). Cox Analysis of Survival Data with Non-Proportional Hazard Functions. *Journal of the Royal Statistical Society, Series D, 41 (2),* 455-465.

Svendsen, G. B., & Prebensen, N. K. (2011). The effect of brand on churn in the telecommunications sector. *European Journal of Marketing, 47 (8),* 1177-1189.

Therneau, T.M., Grambsch, P.M., & Fleming, T.R. (1990). Martinagle-based residuals for survival models. *Biometrika, 77 (1),* 147

The Convergence Consulting Group Limited. (2015). *The Battle for the North American (US/Canada) Couch Potato: Online & Traditional TV and Movie Distribution. Toronto, Ontario.*

Unhanandana, S., & Wattanasupachoke, T. (2012). Customer relationship strategies: the study on customer perspectives. *International Journal of Business and Social Science, 3 (15),* 155-164.

Van den Poel, D., & Leunis, J., 1998. Database marketing modeling for financial services using hazard rate models. *The International Review of Retail, Distribution and Consumer Research 8 (2),* 243–256.

Van den Poel, D. & Larivière, B. (2004). Customer attrition analysis for financial services using proportional hazard models. *European Journal of Operational Research, 157,* 196-217.

Wang, Y., Lo, H-P. & Yang, Y. (2004). An integrated framework for service quality, customer value, satisfaction: evidence from China's telecommunication industry. *Information Systems Frontiers, 6 (4),* 325–340.

Weerahandi, S., & Moitra, S., 1995. Using survey data to predict adoption and switching for services. *Journal of Marketing Research 32 (1),* 85–96.

Wilson, M. G. (2013). Assessing model adequacy in proportional hazards regression. *Statistics and Data Analysis, SAS Global Forum*.

Zhang, H., Lu, Y., Gupta, S., Zhao, L., Chen, A., & Huang, H. (2014). Understanding the antecedents of customer loyalty in the Chinese mobile service industry: a push-pull-mooring framework. *International Journal of Mobile Communications, 12 (6), 551-577*

Zineldin, M. (2006). The royalty of loyalty: CRM, quality and retention. *Journal of Consumer Marketing, 23 (7), 430-437.*

Zeithaml, V. A. (1988). Consumer perception of price, quality, and value: a means-end model and synthesis of evidence. *Journal of Marketing, 52 (3),* 2-22.

Zorn, S., Jarvis, W., & Bellman, S. (2010). Attitudinal perspectives for predicting churn. *Journal of Research in Interactive Marketing, 4 (2),* 157-169.

APPENDIX A

Table 1: Variable summary

	Predictor Variables	Variable name
Usage variables	Basic service usage	LIVE_TV_USG
	Free-VAS usage	ALL_FVAS_USG
	Payable-VAS usage	ALL_PVAS_USG
Subscription variables	Monthly service fee	TOT_REV
	Overdue bill amount	OVERDUE_BILL
	Number of services with the provider	QT_SERVICES
	Overall account subscription duration	SUB_DURATION
	TV subscription duration	TV_SUB_DURATION
	Premium service subscription	PREMIUM_SERV
	Set-top box ownership status	STB_OWN
	Billing credits / promotions	TV_DISCOUNT
Demographic variables	Age	AGE
	Geographical region	REGION
	Type of home	HOME_TYP

Table 2: Detailed variable summary

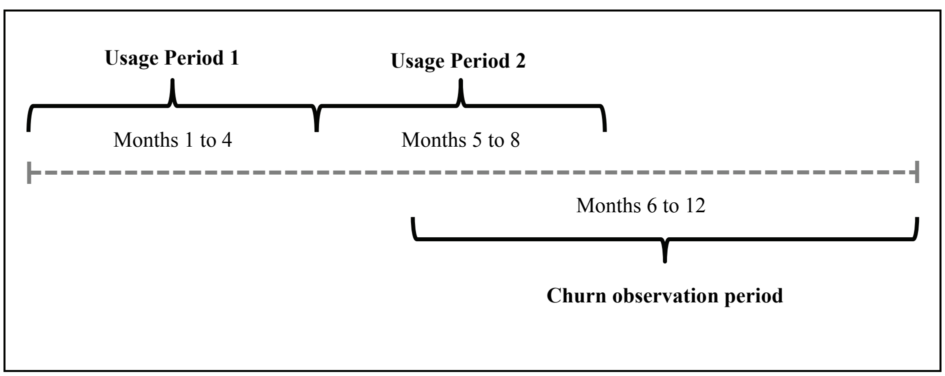
	Description	Variable name	Type of variable
Usage Metrics	Basic service usage	LIVE_TV_USG	Continuous: hours / month
	Free-VAS usage	ALL_FVAS_USG	Continuous: hours / month
	Payable-VAS usage	ALL_PVAS_USG	Continuous: hours / month
Subscription variables	Monthly service fee	TOT_REV	Continuous
	Overdue bill amount	OVERDUE_BILL	Continuous
	Number of other services with the provider	QT_SERVICES	Continuous
	Overall account subscription duration	SUB_DURATION	Continuous
	TV subscription duration	TV_SUB_DURATION	Continuous
	Premium service subscription	PREMIUM_SERV	Categorical: 0: no 1: yes
	Set-top box ownership status	STB_OWN	Categorical: 0: rental 1: purchased 2: purchased & rental
	Billing credits / promotions	TV_DISCOUNT	Continuous
Demographic variables	Age	AGE	Continuous
	Geographical region	REGION	Categorical: 0: urban 1: rural
	Type of home	HOME_TYP	Categorical: 0: apartment 1: house

Table 3: Descriptive statistics

The MEANS Procedure

Variable	*N*	*Mean*	*Std Dev*	*Minimum*	*Maximum*
NO_CLI	137762	51848478.50	20020204.59	10003550.00	75488565.00
DURATION	137762	194.3084886	104.4172721	0	372.0000000
LIVE_TV_USG	137762	156.3130834	127.2766890	0	1288.00
ALL_VAS_USG	137762	2.6857713	9.5023334	0	399.7000000
ALL_FVAS_USG	137762	2.0273350	7.2633164	0	399.7000000
ALL_PVAS_USG	137762	0.6584356	5.3647709	0	235.3000000
TOT_REV	137762	52.5024633	16.3902299	0	152.9100000
OVERDUE_BILL	137762	12.1287587	45.7092798	-12.1000000	1310.65
TV_DISCOUNT	137762	8.2049830	7.5121441	0	60.8400000
QT_SERVICES	137762	2.7877136	0.8240452	1.0000000	4.0000000
STB_OWN	137762	0.3851860	0.6358329	0	2.0000000
PREMIUM_SERV	137762	0.3067464	0.4611449	0	1.0000000
SUB_DURATION	137762	10.4363395	9.8353989	0	56.0000000
TV_SUB_DURATION	137762	4.4899537	3.2818725	0	16.0000000
AGE	137762	55.8423803	16.2352774	18.0000000	103.0000000
REGION	137762	0.9150274	0.2788420	0	1.0000000
HOME_TYP	137762	0.6681596	0.4708757	0	1.0000000
CLUSTER	137762	2.6681959	2.5154126	1.0000000	9.0000000

Figure 1: Research design and observation periods



APPENDIX B

Table 4: Ward's clustering results

Cluster History

Number of Clusters	Clusters Joined		Freq	Semipartial R-Square	R-Square	Approximate Expected R-Square	Cubic Clustering Criterion	Pseudo F Statistic	Pseudo t-Squared	Tie
30	CL78	CL64	212	0.0006	.984	.982	12.5	24E3	297	
29	CL75	CL57	495	0.0006	.983	.981	12.1	24E3	618	
28	CL70	CL71	114	0.0007	.982	.980	11.7	24E3	186	
27	CL50	CL69	287	0.0007	.981	.979	11.2	24E3	327	
26	CL37	CL94	455	0.0008	.981	.979	10.6	24E3	375	
25	CL56	CL65	121	0.0008	.980	.978	10.4	24E3	178	
24	CL32	CL41	311	0.0008	.979	.977	10.4	24E3	179	
23	CL43	CL52	318	0.0011	.978	.976	9.46	23E3	311	
22	CL1278	CL38	6511	0.0012	.977	.975	8.70	23E3	12E3	
21	CL42	CL46	382	0.0012	.976	.974	8.43	23E3	369	
20	CL48	CL68	179	0.0012	.974	.972	8.37	23E3	251	
19	CL20	CL31	379	0.0016	.973	.971	7.46	23E3	182	
18	CL45	CL40	333	0.0017	.971	.969	6.88	23E3	401	
17	CL21	CL30	594	0.0018	.969	.967	6.58	23E3	312	
16	CL29	CL39	767	0.0019	.967	.965	6.60	23E3	802	
15	CL35	CL34	421	0.0021	.965	.963	6.88	23E3	532	
14	CL33	CL26	703	0.0025	.963	.960	6.73	23E3	631	
13	CL25	CL24	432	0.0027	.960	.957	7.19	23E3	318	
12	CL47	CL17	805	0.0034	.957	.954	7.15	23E3	479	
11	CL27	CL23	605	0.0042	.952	.950	6.60	23E3	702	
10	CL22	CL36	7090	0.0045	.948	.944	7.33	24E3	13E3	
9	CL15	CL18	754	0.0057	.942	.938	7.54	24E3	631	
8	CL14	CL28	817	0.0072	.935	.931	7.68	24E3	905	
7	CL8	CL11	1422	0.0115	.923	.921	4.24	23E3	762	
6	CL19	CL9	1133	0.0121	.911	.907	5.29	24E3	785	
5	CL10	CL16	7857	0.0146	.897	.889	9.13	25E3	11E3	
4	CL6	CL13	1565	0.0327	.864	.861	2.77	25E3	1427	
3	CL4	CL12	2370	0.0614	.803	.814	-5.7	24E3	1872	
2	CL5	CL7	9279	0.1090	.694	.687	2.64	26E3	17E3	
1	CL2	CL3	11649	0.6935	.000	.000	0.00	.	26E3	

Figure 2: Ward's cluster selection criteria plots

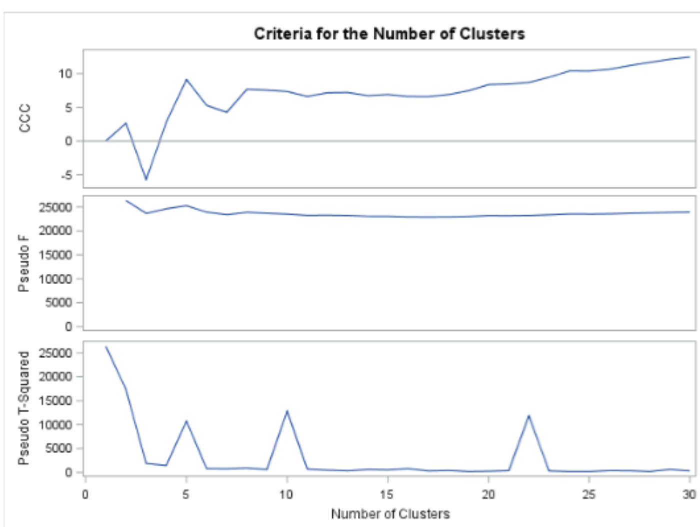


Table 5: VAS user-groups and cluster interpretation

Clusters	Cluster means (1st period)	Cluster means (2nd period)	Type of user	Usage pattern
1	*0.0506*	*0.1206*	*Non-user*	*Constant*
2	*1.6233*	*3.7239*	*Light-Light*	*(+)*
3	*1.7969*	*1.1387*	*Light-Light*	*(-)*
4	*2.0585*	*6.8936*	*Light-Medium*	*(+)*
5	*6.1151*	*7.800*	*Medium-Heavy*	*(+)*
6	*9.0375*	*8.4933*	*Heavy-Heavy*	*(-)*
7	*4.8933*	*1.7566*	*Medium-Light*	*(-)*
8	*7.7284*	*4.9135*	*Medium-Medium*	*(-)*
9	*5.0265*	*5.0053*	*Medium-Medium*	*Constant*

Table 6: K-Means clustering summary

Cluster Summary

Cluster	*Frequency*	*RMS Std Deviation*	*Maximum Distance from Seed to Observation*	*Radius Exceeded*	*Nearest Cluster*	*Distance Between Cluster Centroids*
1	6792	0.3362	1.8801		3	2.0214
2	815	0.9484	2.0648		3	2.5911
3	1305	0.7376	1.6565		1	2.0214
4	188	1.1794	3.7265		2	3.1994
5	365	0.9285	3.0518		9	2.9992
6	667	0.9534	1.8294		5	3.0035
7	534	1.0536	4.4590		3	3.1574
8	416	0.9971	4.1248		9	2.7035
9	567	0.8028	1.4367		8	2.7035

Table 7: K-Means goodness of fit statistics

Statistics for Variables

Variable	*Total STD*	*Within STD*	*R-Square*	*RSQ/(1-RSQ)*
AVG_USG_GR1	2.81944	0.61584	0.952323	19.974561
AVG_USG_GR2	2.71092	0.66194	0.940419	15.783926
OVER-ALL	2.76571	0.63930	0.946605	17.728240

Pseudo F Statistic = 25794.59

Approximate Expected Over-All R-Squared = 0.88914

Cubic Clustering Criterion = 90.670

Table 8: K-Means cluster means and standard deviation

Cluster Means		
Cluster	*AVG_USG_GR1*	*AVG_USG_GR2*
1	0.050647821	0.120583039
2	1.623312883	3.723926380
3	1.796934866	1.138697318
4	2.058510638	6.893617021
5	6.115068493	7.800000000
6	9.037481259	8.493253373
7	4.893258427	1.756554307
8	7.728365385	4.913461538
9	5.026455026	5.005291005

Cluster Standard Deviations		
Cluster	*AVG_USG_GR1*	*AVG_USG_GR2*
1	0.219293638	0.421803508
2	1.102463767	0.763802909
3	0.759998658	0.714602181
4	1.312947419	1.028659739
5	0.948056316	0.908446325
6	0.811019238	1.077144234
7	1.025853171	1.080573869
8	0.881271382	1.100804323
9	0.813537378	0.791947243

Figure 3: Kaplan-Meier survival functions and plots

The LIFETEST Procedure

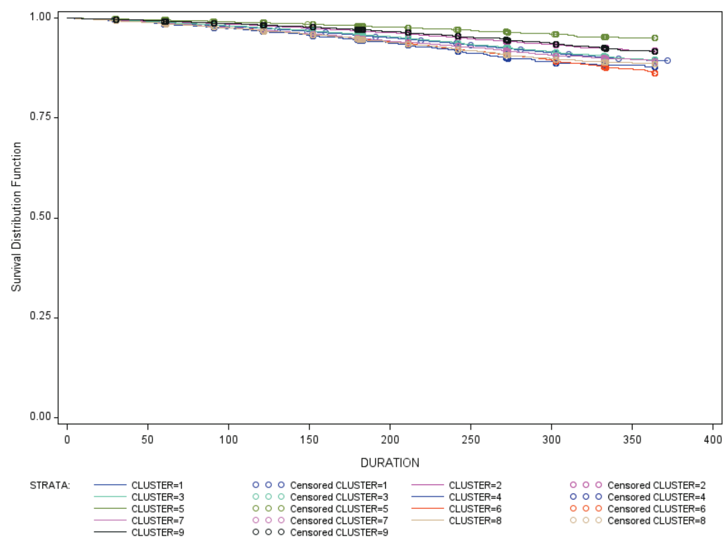


Table 9: Kaplan-Meier survival curve statistical significance

Test of Equality over Strata

Test	*Chi-Square*	*DF*	*Pr > Chi-Square*
Log-Rank	175.2584	8	<.0001
Wilcoxon	159.3712	8	<.0001
-2Log(LR)	191.3895	8	<.0001

Table 10: Cluster censorship distribution

The LIFETEST Procedure

Summary of the Number of Censored and Uncensored Values

Stratum	*CLUSTER*	*Total*	*Failed*	*Censored*	*Percent Censored*
1	1	80293	4305	75988	94.64
2	2	9674	393	9281	95.94
3	3	15452	802	14650	94.81
4	4	2211	149	2062	93.26
5	5	4348	106	4242	97.56
6	6	7883	522	7361	93.38
7	7	6261	366	5895	94.15
8	8	4890	309	4581	93.68
9	9	6750	263	6487	96.10
Total		137762	7215	130547	94.76

APPENDIX D

Table 11: Baseline Cox Proportional Hazard model

Analysis of Maximum Likelihood Estimates							
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
LIVE_TV_USG		1	-0.00207	0.0001158	319.3053	<.0001	0.998
ALL_FVAS_USG		1	-0.0008990	0.00229	0.1542	0.6946	0.999
ALL_PVAS_USG		1	0.00481	0.00209	5.2828	0.0215	1.005
CLUSTER	2	1	-0.40130	0.05333	56.6206	<.0001	0.669
CLUSTER	3	1	-0.01294	0.03907	0.1098	0.7404	0.987
CLUSTER	4	1	0.05966	0.08455	0.4980	0.4804	1.061
CLUSTER	5	1	-0.91828	0.10061	83.3094	<.0001	0.399
CLUSTER	6	1	0.08128	0.06327	1.6504	0.1989	1.085
CLUSTER	7	1	-0.06547	0.05564	1.3844	0.2393	0.937
CLUSTER	8	1	0.06713	0.06286	1.1406	0.2855	1.069
CLUSTER	9	1	-0.32978	0.06477	25.9218	<.0001	0.719
TOT_REV		1	0.00391	0.0008788	19.7895	<.0001	1.004
OVERDUE_BILL		1	0.00327	0.0001555	440.8177	<.0001	1.003
QT_SERVICES	2	1	-0.18106	0.04530	15.9733	<.0001	0.834
QT_SERVICES	3	1	-0.54203	0.04518	143.9429	<.0001	0.582
QT_SERVICES	4	1	-0.79646	0.05500	209.7352	<.0001	0.451
AGE		1	-0.01895	0.0008184	536.3088	<.0001	0.981
SUB_DURATION		1	-0.05157	0.00247	435.7008	<.0001	0.950
TV_SUB_DURATION		1	-0.00105	0.00539	0.0379	0.8456	0.999
PREMIUM_SERV	1	1	-0.17533	0.03127	31.4338	<.0001	0.839
STB_OWN	0	1	0.28823	0.03783	58.0584	<.0001	1.334
STB_OWN	2	1	-0.18308	0.06473	8.0003	0.0047	0.833
TV_DISCOUNT		1	0.01739	0.00157	122.4660	<.0001	1.018
REGION	0	1	-0.16144	0.04806	11.2831	0.0008	0.851
HOME_TYP	0	1	0.17584	0.02549	47.5712	<.0001	1.192

Table 12: Baseline Cox Proportional Hazard model fit statistics

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	162449.35	157656.49
AIC	162449.35	157706.49
SBC	162449.35	157878.59

Figure 4: Baseline model deviance residuals

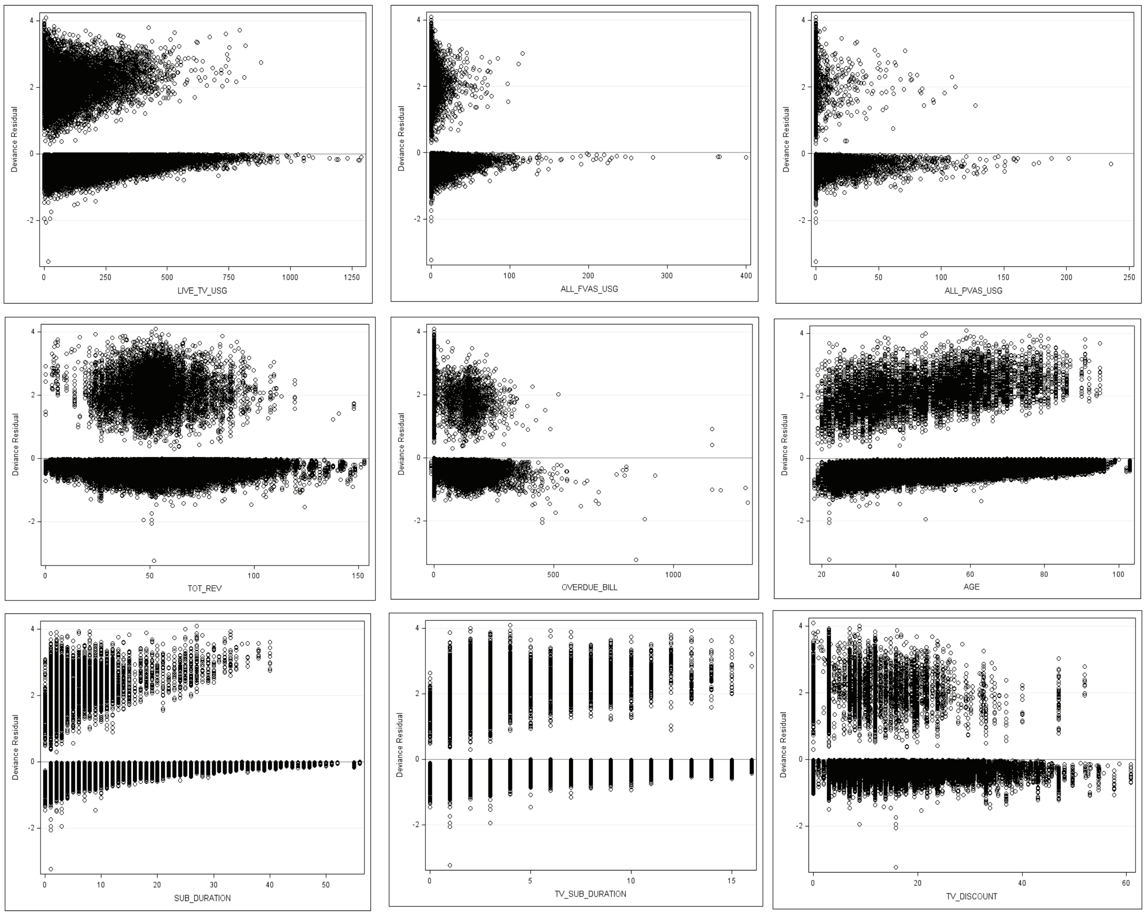


Figure 5: Baseline model DFBETA outlier detection

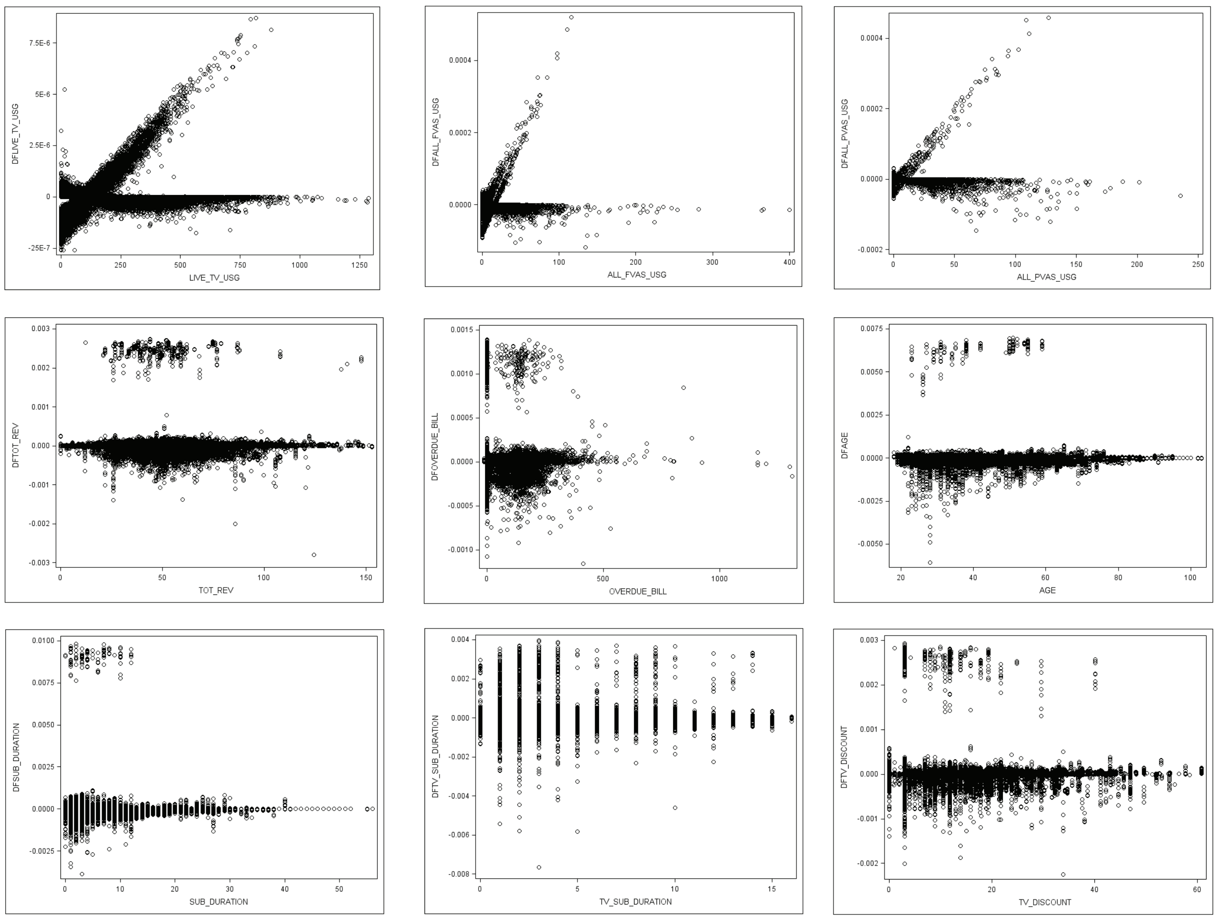
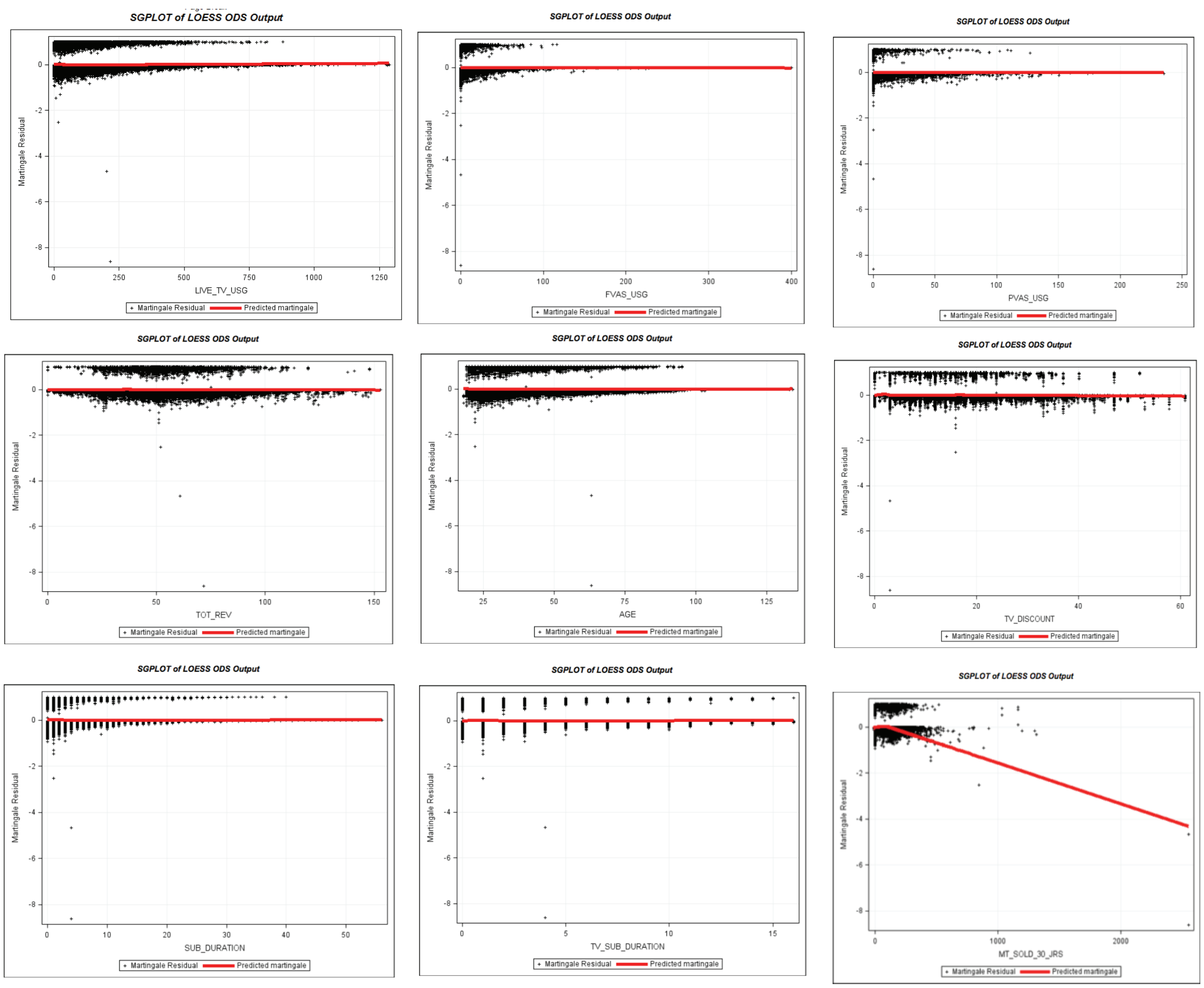


Figure 6: Baseline model functional form



APPENDIX E

Table 13: Log-transformed Cox Proportional Hazard model

Analysis of Maximum Likelihood Estimates							
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
LLIVE_TV_USG		1	-0.20022	0.00715	783.1726	<.0001	0.819
LALL_FVAS_USG		1	-0.01945	0.02207	0.7768	0.3781	0.981
LALL_PVAS_USG		1	0.05149	0.02481	4.3052	0.0380	1.053
CLUSTER	2	1	-0.39108	0.05437	51.7468	<.0001	0.676
CLUSTER	3	1	-0.00580	0.03957	0.0215	0.8835	0.994
CLUSTER	4	1	0.09617	0.08644	1.2377	0.2659	1.101
CLUSTER	5	1	-0.85977	0.10603	65.7476	<.0001	0.423
CLUSTER	6	1	0.13659	0.07650	3.1878	0.0742	1.146
CLUSTER	7	1	-0.04830	0.05835	0.6853	0.4078	0.953
CLUSTER	8	1	0.10779	0.07073	2.3223	0.1275	1.114
CLUSTER	9	1	-0.27654	0.06912	16.0085	<.0001	0.758
LTOT_REV		1	0.17768	0.03608	24.2490	<.0001	1.194
LOVERDUE_BILL		1	0.13263	0.00640	429.7650	<.0001	1.142
QT_SERVICES	2	1	-0.31366	0.04832	42.1342	<.0001	0.731
QT_SERVICES	3	1	-0.63063	0.04936	163.2574	<.0001	0.532
QT_SERVICES	4	1	-0.83465	0.05869	202.2358	<.0001	0.434
LAGE		1	-0.82366	0.03720	490.2644	<.0001	0.439
LTV_SUB_DURATION		1	0.22214	0.03276	45.9799	<.0001	1.249
LSUB_DURATION		1	-0.56526	0.02442	535.8001	<.0001	0.568
PREMIUM_SERV	1	1	-0.19144	0.02965	41.7006	<.0001	0.826
STB_OWN	0	1	0.22390	0.03808	34.5733	<.0001	1.251
STB_OWN	2	1	-0.25741	0.06484	15.7580	<.0001	0.773
LTV_DISCOUNT		1	0.19347	0.01736	124.2597	<.0001	1.213
REGION	0	1	-0.18012	0.04803	14.0626	0.0002	0.835
HOME_TYP	0	1	0.14991	0.02549	34.6023	<.0001	1.162

Table 14: Log-transformed model fit statistics

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	130323.57	124981.42
AIC	130323.57	125031.42
SBC	130323.57	125203.52

Table 15: Log-transformed model – Proportional hazards test

Pearson Correlation Coefficients, N = 7215
Prob > |r| under H0: Rho=0

	rLLIVE_TV_USG	rLALL_FVAS_USG	rLALL_PVAS_USG	rCLUSTER	rLTOT_REV	rLOVERDUE_BILL	rQT_SERVICES
TIMERANK Rank for Variable DURATION	0.16655 <.0001	0.05699 <.0001	-0.01085 0.3568	-0.00047 0.9683	-0.00549 0.6413	-0.00337 0.7749	0.01618 0.1694

rLAGE	rLTV_SUB_DURATION	rLSUB_DURATION	rPREMIUM_SERV	rSTB_OWN	rLTV_DISCOUNT	rREGION	rHOME_TYP
-0.01764	-0.00989	0.01499	0.01736	0.05182	-0.00366	0.04454	0.03627
0.1341	0.4009	0.2031	0.1403	<.0001	0.7561	0.0002	0.0021

Figure 7: Log-transformed model functional form

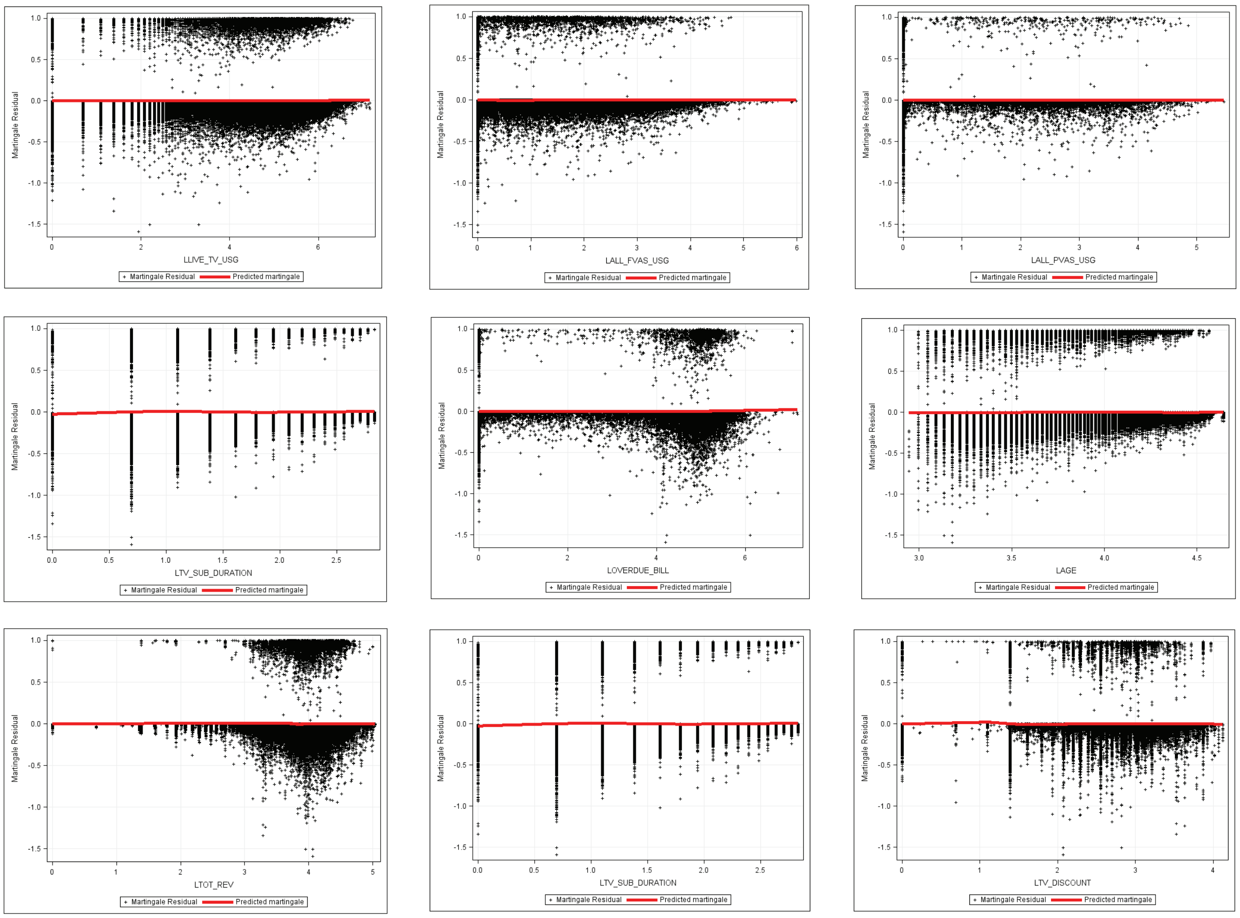
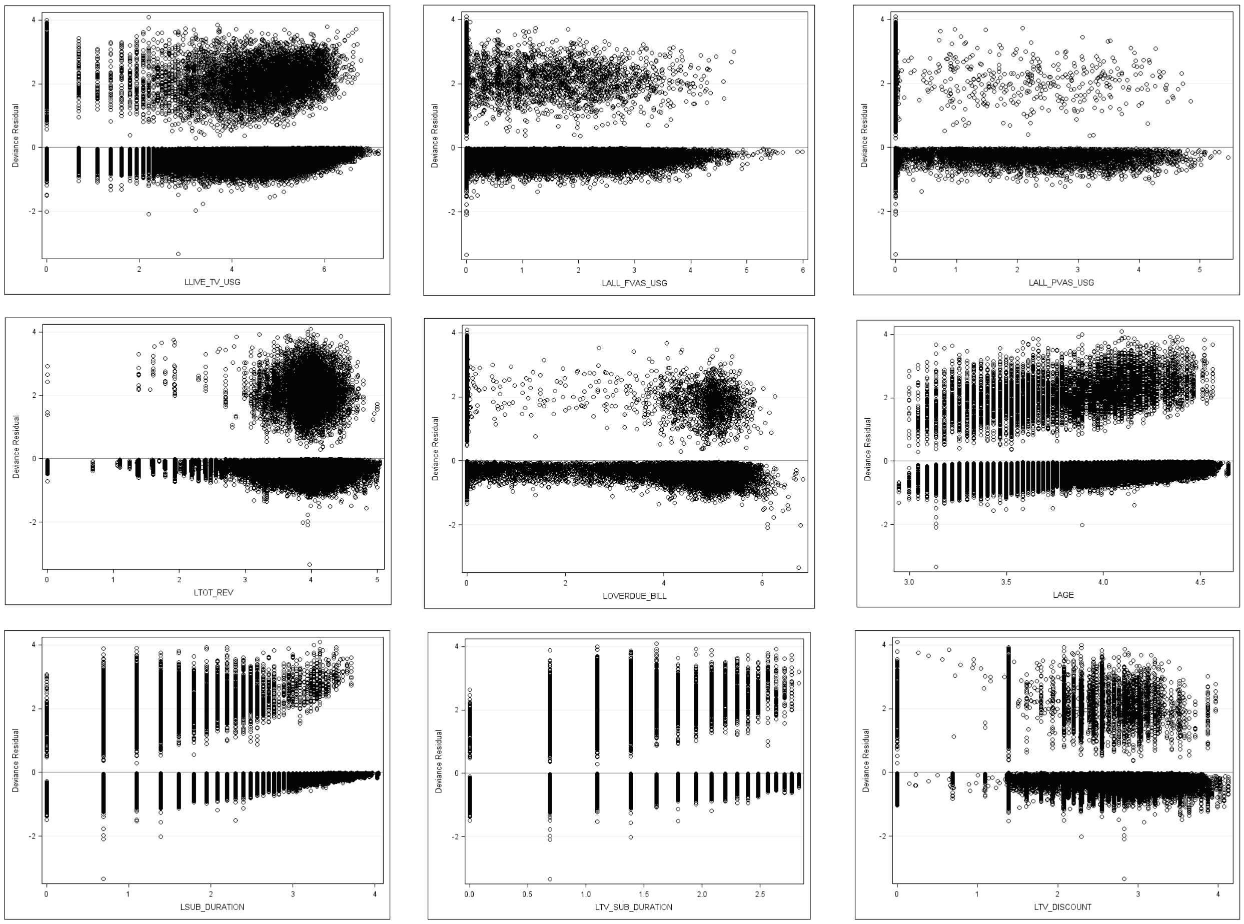


Figure 8: Log-transformed model deviance residuals



APPENDIX F

Table 16: Final Cox Proportional Hazard model fit statistics

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	130323.57	124763.38
AIC	130323.57	124817.38
SBC	130323.57	125003.25

Table 17: Final Cox Proportional Hazard model significance levels

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
LLIVE_TV_USG	1	762.6040	<.0001
LALL_FVAS_USG	1	7.0425	0.0080
LALL_PVAS_USG	1	5.4749	0.0193
CLUSTER	8	158.8872	<.0001
LTOT_REV	1	23.1731	<.0001
LOVERDUE_BILL	1	422.7163	<.0001
QT_SERVICES	3	296.3553	<.0001
LAGE	1	509.9221	<.0001
LTV_SUB_DURATION	1	46.4681	<.0001
LSUB_DURATION	1	535.2667	<.0001
LTV_DISCOUNT	1	121.7437	<.0001
PREMIUM_SERV	1	41.4384	<.0001
HOME_TYP	1	32.7382	<.0001
STB_OWN	2	100.6646	<.0001
REGION	1	15.3765	<.0001
itLLIVE_TV_USG	1	188.7891	<.0001
itLALL_FVAS_USG	1	8.7035	0.0032

Table 18: Final Cox Proportional Hazard model covariate elimination

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	itSTB_OWN	1	19	0.0298	0.8630
2	itHOME_TYP	1	18	0.1758	0.6750
3	itREGION	1	17	0.5044	0.4776

Table 19: Final Cox Proportional Hazard model

Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
LLIVE_TV_USG		1	-0.35320	0.01279	762.6040	<.0001	0.702
LALL_FVAS_USG		1	-0.08697	0.03277	7.0425	0.0080	0.917
LALL_PVAS_USG		1	0.05790	0.02475	5.4749	0.0193	1.060
CLUSTER	2	1	-0.38170	0.05437	49.2911	<.0001	0.683
CLUSTER	3	1	-0.01097	0.03958	0.0769	0.7816	0.989
CLUSTER	4	1	0.10697	0.08642	1.5319	0.2158	1.113
CLUSTER	5	1	-0.86374	0.10599	66.4084	<.0001	0.422
CLUSTER	6	1	0.11329	0.07662	2.1862	0.1392	1.120
CLUSTER	7	1	-0.06163	0.05841	1.1133	0.2914	0.940
CLUSTER	8	1	0.08236	0.07093	1.3480	0.2456	1.086
CLUSTER	9	1	-0.28609	0.06916	17.1114	<.0001	0.751
LTOT_REV		1	0.17341	0.03602	23.1731	<.0001	1.189
LOVERDUE_BILL		1	0.13156	0.00640	422.7163	<.0001	1.141
QT_SERVICES	2	1	-0.31398	0.04832	42.2306	<.0001	0.731
QT_SERVICES	3	1	-0.62891	0.04934	162.4446	<.0001	0.533
QT_SERVICES	4	1	-0.83430	0.05869	202.1018	<.0001	0.434
LAGE		1	-0.84021	0.03721	509.9221	<.0001	0.432
LTV_SUB_DURATION		1	0.22347	0.03278	46.4681	<.0001	1.250
LSUB_DURATION		1	-0.56478	0.02441	535.2667	<.0001	0.568
LTV_DISCOUNT		1	0.19153	0.01736	121.7437	<.0001	1.211
PREMIUM_SERV	1	1	-0.19084	0.02965	41.4384	<.0001	0.826
HOME_TYP	0	1	0.14587	0.02549	32.7382	<.0001	1.157
STB_OWN	0	1	0.22143	0.03810	33.7836	<.0001	1.248
STB_OWN	2	1	-0.26656	0.06484	16.8987	<.0001	0.766
REGION	0	1	-0.18836	0.04803	15.3765	<.0001	0.828
itLLIVE_TV_USG		1	0.00119	0.0000866	188.7891	<.0001	1.001
itLALL_FVAS_USG		1	0.0004447	0.0001507	8.7035	0.0032	1.000

Table 20: Results and Hypothesis Summary

Hypotheses		Results
H1	Basic (core) service usage has a negative influence on customer churn.	*Supported*
H2a	VAS users are less likely to churn than non-users.	*Supported*
H2b	Free VAS usage has a greater influence on reducing churn susceptibility than payable VAS usage.	*Supported*
H3a	VAS users with greater overall VAS usage frequency are less likely to churn than those with lower usage frequency.	*Partially supported*
H3b	VAS users with increasing VAS usage pattern are less likely to churn than VAS users with decreasing usage pattern	*Not Supported*
H3c	Heaviest VAS users exhibit the greatest risk of churning	*Supported*