American Journal of Theoretical and Applied Statistics
2015; 4(6): 471-479
Published online October 12, 2015 (http://www.sciencepublishinggroup.com/j/ajtas)
doi: 10.11648/j.ajtas.20150406.17
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

Extended Cox Modeling of Customer Retention in Mobile Telecommunication Sector of Rwanda

Diane Ingabire*, Samuel Musili Mwalili, George Otieno Orwa

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email address:
ingadiane@gmail.com (D. Ingabire)

To cite this article:
Diane Ingabire, Samuel Musili Mwalili, George Otieno Orwa. Extended Cox Modeling of Customer Retention in Mobile Telecommunication Sector of Rwanda. *American Journal of Theoretical and Applied Statistics.* Vol. 4, No. 6, 2015, pp. 471-479.
doi: 10.11648/j.ajtas.20150406.17

Abstract: Retaining customers improves profitability, importantly reduces the cost incurred in acquiring new customers and moreover a firm can increase profits by 25-95 percent if it could improve its customer retention rates by 5 percent. As markets mature and competitive pressure intensifies, companies can no longer ignore the importance of customer retention as their existing customer bases have become their precious assets. This research aims to model customer retention in Rwandan telecom sector using survival analysis technique in order to inform the concerned institutions and companies about telecom customer retention in Rwanda. The Cox regression model and extended Cox model were developed using simulation approach in order to assess which model is the best for customer retention. It was found that the customer's socio-economic, demographic and behavioral characteristics have an effect on churn rate. The extended Cox model was the best description of how customer retention is achieved. These findings hold implications for industry operators on key areas to pay attention to in order to achieve customer retention.

Keywords: Customer retention, Cox model, Extended Cox Model

1. Introduction

Mobile communication has become the backbone of the society. The mobile telecommunication sector continues to offer unprecedented opportunities to economic growth in both developing and developed markets and mobile services have become an essential part of how economy works and functions. The total mobile penetration has more than doubled in all the region of the world since 2005(Williams et al. 2012).

Customer retention refers to customer's stated continuation of a business relationship with the firm (Timothy et al. 2007). Since last decade, many companies perceive the retention of the customer as a central topic in their management and marketing decisions (van Den Poel & Lariviere 2004). Many studies on customer retention argue that retaining customers improves profitability, importantly reduces the cost incurred in acquiring new customers. Reichheld & Schefter (2000) discovered that a firm can increase profits by 25-95 percent if it could improve its customer retention rates by 5 percent. Furthermore, a retained customer will be loyal due to the attachment and commitment to the organization (Sharmeela-Banu et al. 2012).

As markets mature and competitive pressure intensifies, companies can no longer ignore the importance of customer retention as their existing customer bases have become their precious assets. Customer churn is a focal concern for most companies which are active in industries with low switching cost and among all industries that suffers from this issue, telecommunications industries can be considered at the top of the list with approximate annual churn rate of 30% (Ali, 2009).

Rwanda is a poor, small and landlocked country with surface area of 26,338 square kilometers situated in east-central Africa. In Rwanda, regular efforts have been made to develop the service sector and to stimulate investment in the industrial sector and Vision 2020 seeks to transform Rwanda from a low-income agriculture-based economy to a knowledge-based, service-oriented economy with a middle-income country status by 2020 and ICT is one of the cross cutting issues of vision 2020 (RDB, 2013).

The Rwandan telecom sector has shown particularly strong growth in recent years in terms of subscriptions, revenues and investments, buttressed by a vibrant economy and a GDP which has sustained growth of between 7% and 8% annually

since 2008. As a result, the country is rapidly catching up with other markets in Africa, with increased penetration particularly evident in the internet and mobile sectors. Although the country was slow to liberalize the mobile sector, there is effective competition among the three current operators, each of which provides wide geographic coverage (Dudde, 2014). The deregulation of the industry has caused a lots of service providers to enter the industry and it can be stated that the telecommunication industry has been very competitive; as at the end of December 2014, there were two fixed line telephony operators and three mobile telephony operators who are fully operational, namely, Mobile Telecommunication Network (MTN) having both fixed line telephony and mobile telephony, Millicom Rwanda Limited (Tigo), Airtel Rwanda and Rwandatel (RURA, 2015).

As the competition continues to increase in the telecommunications industry, customers are continuously leaving one company to another and retaining customers has become a critical concern for most companies; hence the cellular phone companies are doing everything possible to attract new customers and retain the existing ones. However, despite the large amount of research done on customer retention in mobile telecommunications, there is an absence of studies about the sector in Rwanda. In order to support telecommunications companies manage churn reduction, not only do we need to predict which customers are at high risk of churn, but also we need to know how soon these high-risk customers will churn; so that the telecommunications companies can optimize their marketing intervention resources to prevent as many customers as possible from churning. Conventional statistical methods (e.g. logistics regression, decision tree, and etc.) are very successful in predicting customer churn, but these methods could hardly predict when customers will churn, or how long the customers will stay with. However, survival analysis was, at the beginning, designed to handle survival data, and therefore is an efficient and powerful tool to predict customer churn (Lu, 2002).

Survival analysis is a collection of statistical procedures for data analysis for which the outcome variable of interest is time until an event occurs. Survival analysis examines and models the time it takes for events to occur. Survival modeling examines the relationship between survival and one or more predictors (covariates). The main aim of this paper is to model customer retention in telecommunication sector using survival analysis technique in order to inform the concerned institutions and companies about telecom customer retention in Rwanda.

2. Review of Previous Research

Many studies have been done on customer retention in telecommunication sector. In the study aimed to understand and predict customer lifetime in a contractual setting in order to improve the practice of customer portfolio management, Portela & Menezes (2011) used Accelerated Failure Time (AFT) models after estimating the Cox PH model in order to test the PH assumption based on Schoenfeld residuals and found the PH assumption did not hold. Appropriate parametric model was found to be a log-logistic based on the fact that it had the lowest AIC. Wong (2011) instead used the Cox regression model in studying customer retention in the context of a Canadian wireless telecommunications company and explored the predictors of churn incidence as part of customer relationship management. On the other hand, Ahn et al. (2006) also investigated factors leading to customer churn using a sample of 5789 actual customer transactions and billing data. In addition, the mediating effects of customer status between churn determinants and customer churn were analyzed by adopting logistic regressions. The parameters of logistic response functions were estimated with the maximum likelihood method. The likelihood ratio test indicates that these models fit the data very well but this study had a number of limitations in that: first, data for some variables, such as account tenure (also called customer duration) and each subscriber's age were not available and in particular the account tenure is a very important variable explaining customer churn. Secondly, the 8 months data collection period for the study was relatively short which suggested that an additional longitudinal study with a longer period of data collection and time-series data was necessary. Ocloo & Tsetse (2013) instead adopted the descriptive survey method which employed a hybrid of qualitative and quantitative methods in their study on customer retention in the Ghanaian mobile telecommunication industry.

It is evident that many models has been used to model telecom customer retention globally; however, all the models reviewed above have a number of weakness which make them less suitable than extended Cox regression adopted in this research. The logistic regression models the outcome as a binary variable taking value 1 or 0, it hence ignores the effect of survival times and censoring information. The AFT models make assumptions on the distributions of the survival times and does not take into account the effect of time-dependent variables; Cox regression model ignores the effect of time dependent variables and the other customer retention models used are weak compared to extended Cox model in the fact that they did not take into account the effect of survival times and censoring information; thus the main aim of this paper is to model telecom customer retention using Cox regression model and extended Cox model and determine which is the best retention model.

3. Methodology

3.1. Basic Analytical Quantities

Let T represents the survival time; the actual survival time of an individual t can be regarded as the value of the variable T which can take any non- negative value. T is regarded as random variable with cumulative distribution function $F(t) = Pr\ (T \leq t)$ and probability density function $f(t) = dF(t)/dt$. The basic analytical quantities for time-to-event data are the survival function

$$S(t) = Pr\ (T > t) = 1 - F(t) \quad (1)$$

which gives the probability that a customer survives longer than some specified time t and the hazard function

$$h(t) = lim_{\delta t \to 0} \left\{ \frac{Pr(t \le T < \delta t | T \ge t)}{\delta t} \right\} = \frac{f(t)}{S(t)} \quad (2)$$

which is also referred to as hazard rate , instantaneous failure rate, or conditional failure rate. In the context of this study, it can be interpreted as the risk of canceling the contract (or unsubscribing) at time t. Another basic quantity is the cumulative hazard function defined as

$$H(t) = \int_0^t h(u)du = -logS(t) \quad (3)$$

3.2. Cox Model

A Cox model is a statistical technique for exploring the relationship between the survival of an individual and several explanatory variables. It allows us to estimate the hazard (or risk) of death (or cancelation) for an individual given their prognostic variables. The general proportional model is given by

$$h_i(t) = exp\big(\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots \beta_p x_{pi}\big) h_0(t) = h_0(t) e^{\beta' x_i} \quad (4)$$

Where x_i is a set of covariates for the i^{th} individual and $h_0(t)$ is the baseline hazard function that depends only on the time and not on the covariates.

The hazard is a product of two terms: the baseline hazard function and the set of covariates which does not depend on time t. The most important assumption in Cox model is the proportional hazards assumption; that is, the hazard ratio of any two individuals is constant over time in the setting where the predictor variables do not vary over time.

In estimating the parameters in the Cox regression model, Cox (1972) derived the same likelihood, and generalized it for censoring, using the idea of a partial likelihood.

Suppose we observe (t_i, δ_i, X_i) for customer i, where t_i is survival times, δ_i is the failure (cancelation)/ censoring indicator (1=cancel, 0=censor), X_i represents a set of covariates and i = 1, 2,.....n. Then the likelihood function is given by

$$L(\beta) = \prod_{i=1}^{n_i} \left[\frac{exp(\beta x_i)}{\sum_{j \in R(t_i)} exp(\beta x_j)} \right]^{\delta_i} \quad (5)$$

Where $R(t_i)$ is the risk set at time t_i. The corresponding log-likelihood function is given by

$$nlogL(\beta) = \sum_{j=1}^{n} \delta_i \left[\beta x_j - log \left[\sum_{l \in R_{(\tau_j)}} \exp(\beta x_l) \right] \right] \quad (6)$$

The maximum likelihood estimates of β-parameters can be found by maximizing this log-likelihood function using numerical methods.

The Newton-Raphson procedure is a numerical method used to fit models for censored survival data by maximizing the partial likelihood function.

Let $u(\beta)$ be the $p \times 1$ vector of the first derivatives of the log-likelihood function in equation (3.6) with respect to the β-parameters. This quantity is known as the vector of efficient score. Also let $I(\beta)$ be the $p \times p$ matrix of negative second derivatives of the log-likelihood, so that the $(j,k)^{th}$ element of $I(\beta)$ is

$$-\frac{\partial^2 logL(\beta)}{\partial \beta_j \beta_k} \quad (7)$$

The matrix $I(\beta)$ is known as the observed information matrix.

According to the Newton-Raphson procedure, an estimate of the vector of β −parameters at the $(s + 1)^{th}$ cycle of the iterative procedure is $\hat{\beta}_{s+1}$ and is given by

$$\hat{\beta}_{s+1} = \hat{\beta}_s + I^{-1}\left(\hat{\beta}_s\right) u\left(\hat{\beta}_s\right) \quad (8)$$

for s = 0,1,2,....., where $u(\hat{\beta}_s)$ is the vector of the efficient score and $I^{-1}(\hat{\beta}_s)$ is the inverse of the information matrix, both evaluated at β_s. The procedure can be started by taking β_s= 0. The process is terminated when the change in the log-likelihood function is sufficiently small or when the largest of relative changes in the values of parameters estimates is sufficient small.

3.3. The Extended Cox Model

One of the strengths of the Cox model is its ability to encompass covariates that change over time; such covariates are known as time-dependent variables. If time-dependent variables are considered, the Cox model form may still be used, but such a model no longer satisfies the PH assumption and is called the extended Cox model. The extended Cox model for time-dependent variables is given

$$h(t, X) = h_0(t) exp\left[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{j=1}^{p2} \delta_j X_j(t)\right] \quad (9)$$

where $X(t) = \{X_1, X_2, \dots \dots X_{p1}; X_1, X_2, \dots \dots X_{p2}\}$ is the entire collection of predictors at time t.

The extended model contains a baseline hazards function $h_0(t)$ and an exponential function which contains both time-independent predictors, as denoted by the X_i variables, and time-dependent predictors, as denoted by the $X_j(t)$ variables. Even though the values of the variable $X_j(t)$ may change over time, the hazard model provides only one coefficient for each time-dependent variable in the model. Thus, at time t, there is only one value of the variable $X_j(t)$ that has an effect on the hazard, that value being measured at time t.

An important assumption of the extended Cox model is that the effect of a time-dependent variable $X_j(t)$ on the survival probability at time t depends on the value of this variable at that same time t, and not on the value at an earlier or later time.

The PH assumption is not satisfied for the extended Cox model, thus the hazard ratio depend on time t. The hazard ratio for the extended Cox model is then

$$\widehat{HR}(t) = exp\left[\sum_{i=1}^{p_1} \beta_i [X_i^* - X_i] + \sum_{j=1}^{p_2} \delta_j \left[X_j^*(t) - X_i(t)\right]\right] \quad (10)$$

The two sets of predictors, $X^*(t)$ and $X(t)$, identify two

specifications at time t for the combined set of predictors containing both time-independent and time-dependent variables.

The extended Cox model for customer retention used in this study is given by

$$h(t,X) = h_0(t)\exp[\sum_{i=1}^{6}\beta_i X_i + \alpha_7 x_7] \quad (11)$$

where $h_0(t)$ is a baseline hazard function

$X_i = \{x_1, x_1x_2, x_3, x_4, x_5, x_6\}$ is a set of time independent covariates where,

x_1 =sex, x_2 =age, x_3 = marital status, x_4 =residence, x_5=employment status, x_6= number of subscriptions

x_7=plan rate and it is the time-varying covariate

β_i= set of coefficients of time independent covariates and α_7=the coefficient of time-varying covariate.

The regression coefficients in the extended Cox model are estimated using a maximum likelihood (ML) procedure. ML estimates are obtained by maximizing a (partial) likelihood function L. The partial log likelihood function from equation (3.6) can be generalized to the case of extended Cox model, thus, the partial log-likelihood becomes

$$\sum_{i=1}^{n}\delta_i\left\{\sum_{j=1}^{p}\beta_j x_{ji}(t_i) - log\sum_{l\in R(t_i)}\exp\left(\sum_{j=1}^{p}\beta_j x_{jl}(t_i)\right)\right\} \quad (12)$$

where $R(t_i)$ is the risk set at time t_i, the death time of the i^{th} individual in the study, i = 1,2,.....,n and δ_i is an event indicator that is zero if the survival time is censored and unity otherwise. The estimates of β−parameters are obtained by maximizing the partial log-likelihood function.

3.4. Simulation to Assess the Validity

R statistical software (version 3.0.3) was used for simulating and analyzing data. Due to unavailability of data and inconsistent recording of telecom data, simulation approach has been used to generate data and customer retention model. According to different research, some customer's socio-economics, demographics and behavioral characteristics have been found to affect the hazard of churning. In this view, six time independent covariates namely the customer's sex, age, marital status, residence, employment status and number of subscriptions and one time dependent covariate namely customer's plan rate were simulated. The maximum following time was set to be 25 months and furthermore permalgorithm function was used to generate survival times condition to this set of covariates and coxph was used to fit the Cox model. In modeling telecom customer retention two models were fitted: the correct model and incorrect model in order to determine which model explains well telecom customer retention. The correct model is the extended Cox model which includes the time dependent variable and the incorrect model is the Cox regression model with time independent variables only, that is; it assumed the time dependent variable (TDV) "plan" to be time independent variable (TIV). This means that dataset1 generated to fit a correct model contained seven covariates in which "plan rate" covariate is time varying covariate and dataset2 generated to fit an incorrect assumed plan rate covariate to be constant over time. The appendices A and B show the first fifty observations generated in the two data sets for both models. The validity of the models were tested through simulation where 500 and 1000 simulations were run for both models and two samples of size 500 and 1000 have been generated for both models in order to assess the effect of sample size in both models.

4. Results and Discussion

Table 4.1. *Summary of estimates of coefficients of covariates in Cox regression model (incorrect model).*

R \ n	500						1000					
	$\hat{\beta}$	coef*	exp(coef*)	SE#	Bias	Pvalues	Freq(p<0.05)	coef*	exp(coef*)	SE#	Bias	Pvalues	Freq(p<0.05)
500	$\hat{\beta}_1$	0.6732	1.9605	0.2734	-0.0199	0.0859	346	0.6728	1.9598	0.1885	-0.0203	0.0113	479
	$\hat{\beta}_2$	-0.2783	0.7570	0.0298	0.0040	0.0000	500	-0.2751	0.7595	0.0206	0.0073	0.0000	500
	$\hat{\beta}_3$	-0.9974	0.3688	0.2657	0.0354	0.0107	475	-1.0014	0.3674	0.1831	0.0314	0.0001	500
	$\hat{\beta}_4$	-0.6236	0.5360	0.2663	0.0112	0.1010	320	-0.6108	0.5429	0.1838	0.0241	0.0211	456
	$\hat{\beta}_5$	-0.7653	0.4652	0.2625	0.0178	0.0418	419	-0.7623	0.4666	0.1809	0.0208	0.0038	491
	$\hat{\beta}_6$	-0.7053	0.4940	0.2322	0.0081	0.0338	431	-0.6988	0.4972	0.1590	0.0146	0.0025	495
	$\hat{\beta}_7$	-0.0360	0.9646	0.1360	0.9211	0.4732	30	-0.0438	0.9572	0.0935	0.9133	0.4595	34
1000	$\hat{\beta}_1$	0.6826	1.9791	0.2731	-0.0105	0.0740	724	0.6771	1.9681	0.1877	-0.0161	0.0106	948
	$\hat{\beta}_2$	-0.2810	0.7551	0.0300	0.0014	0.0000	1000	-0.2784	0.7570	0.0208	0.0039	0.0000	1000
	$\hat{\beta}_3$	-1.0019	0.3672	0.2651	0.0309	0.0067	968	-1.0131	0.3631	0.1826	0.0197	0.0001	1000
	$\hat{\beta}_4$	-0.6222	0.5368	0.2655	0.0127	0.0924	653	-0.6187	0.5387	0.1830	0.0162	0.0172	927
	$\hat{\beta}_5$	-0.7617	0.4669	0.2623	0.0214	0.0423	825	-0.7526	0.4711	0.1801	0.0305	0.0032	986
	$\hat{\beta}_6$	-0.6974	0.4979	0.2308	0.0159	0.0322	855	-0.6934	0.4999	0.1585	0.0199	0.0022	988
	$\hat{\beta}_7$	-0.0442	0.9567	0.1360	0.9129	0.4673	77	-0.0465	0.9546	0.0935	0.9106	0.4600	98

* Average value of the covariates' coefficients estimates, # average standard errors.

Table 4.2. *Summary of estimates of coefficients of covariates in Cox model with time varying covariate (correct model).*

R \ n		500						1000					
	$\hat{\beta}$	coef*	exp(coef*)	SE#	Bias	Pvalues	Freq(p<0.05)	coef*	exp(coef*)	SE#	Bias	Pvalues	Freq(p<0.05)
500	$\hat{\beta}_1$	0.6903	1.9944	0.2742	-0.0028	0.0731	356	0.6946	2.0029	0.1892	0.0014	0.0090	486
	$\hat{\beta}_2$	-0.2844	0.7525	0.0303	-0.0020	0.0000	500	-0.2810	0.7550	0.0209	0.0013	0.0000	500
	$\hat{\beta}_3$	-1.0252	0.3587	0.2672	0.0076	0.0091	478	-1.0296	0.3572	0.1837	0.0033	0.0000	500
	$\hat{\beta}_4$	-0.6366	0.5291	0.2673	-0.0017	0.0914	326	-0.6240	0.5358	0.1843	0.0109	0.0161	467
	$\hat{\beta}_5$	-0.7865	0.4555	0.2633	-0.0034	0.0364	423	-0.7820	0.4575	0.1814	0.001	0.0030	491
	$\hat{\beta}_6$	-0.7198	0.4869	0.2331	-0.0064	0.0278	446	-0.7155	0.4889	0.1594	-0.0021	0.0024	494
	$\hat{\beta}_7$	-0.9481	0.3875	0.1546	0.0090	0.0000	500	-0.9511	0.3863	0.1081	0.0061	0.0000	500
1000	$\hat{\beta}_1$	0.6997	2.0131	0.2741	0.0065	0.0690	733	0.6934	2.0004	0.1882	0.0002	0.0083	958
	$\hat{\beta}_2$	-0.2869	0.7506	0.0305	-0.0045	0.0000	1000	-0.2843	0.7526	0.0210	-0.0019	0.0000	1000
	$\hat{\beta}_3$	-1.0261	0.3584	0.2661	0.0067	0.0053	977	-1.0385	0.3540	0.1832	-0.0057	0.0000	1000
	$\hat{\beta}_4$	-0.6423	0.5261	0.2664	-0.0074	0.0804	682	-0.6339	0.5305	0.1835	0.0010	0.0143	938
	$\hat{\beta}_5$	-0.7845	0.4563	0.2634	-0.0014	0.0344	859	-0.7738	0.4613	0.1807	0.0093	0.0023	989
	$\hat{\beta}_6$	-0.7168	0.4883	0.2317	-0.0035	0.0262	870	-0.7126	0.4904	0.1591	0.0007	0.0014	995
	$\hat{\beta}_7$	-0.9401	0.3906	0.1537	0.0170	0.0000	1000	-0.9468	0.3880	0.1075	0.0103	0.0000	1000

Table 4.1 and table 4.2 show the incorrect and correct model respectively. In both models, the first column represents the average estimated value of coefficients of each covariate, the estimates of coefficients of time independent covariates are mostly equal in both models but the estimates of coefficients of time dependent covariate "plan" are different in both models. The exponentiated coefficients in the second column of the table are interpretable as multiplicative effects on the hazard. For example, taking n=1000 and R=1000 and holding other covariates constant in the correct model, an additional year of age reduces the hazard of churning by a factor of 0.7526 or 24.74 percent and an additional in customer's number of subscriptions reduces the hazard of churning by a factor of 0.4904, or 50.96 percent. The third column represents the average standard errors for each covariate. The standard error of the sample is an estimate of how far the sample mean is likely to be from the population mean and it tends to zero with the increasing sample size. This is also true for our research as the standard errors tend to decrease with the increase of the sample size. The fourth column represents the bias of the estimator, which is the difference between the estimator's expected value and the true value of the parameter being estimated, was determined and it was found that the bias of the estimate of coefficient of time varying covariate in the incorrect model is larger than bias of the estimate of coefficient of time varying covariate in the correct model. Also the average probability values and proportion of p-values which are significant for each covariates have been determined, and the covariate "plan" is significant in the correct but not in the incorrect model. Thus, by ignoring the effect of time dependent variable we can reject the variable plan when we don't have to reject it.

Furthermore in assessing the effect of the sample size on variables' significance, it was found that the variables become more significant with the increase of the sample size, thus all variables in the correct model were found to significantly affect the risk of churning highly for n=1000.

All the estimated values for the six TIV in both models are almost the same. A great difference lies between the estimated values of the variable "plan" for two models. Table 4.3 shows a comparison between the two models for the estimates of coefficient of time varying covariate "plan". It shows the bias, the standard error and the mean squared errors (MSE) for the two models each with two samples of size 500 and 1000. The bias shows how the estimate is close to the true value and the standard error shows how far the sample mean is likely to be from the population mean. The smaller the bias and standard error, the better is the estimate but it is common to trade-off some increase in bias for a larger decrease in the standard error and vice-versa. The bias and the standard error decrease as the sample size increases. The mean squared error captures the error that the estimator makes. The smaller the MSE the better the estimate is. The results from the table below shows that the MSE of the estimate of the coefficient for the correct model is smaller than the MSE for the incorrect model; thus, ignoring the effect of time varying covariate increases the MSE in the incorrect model and this means the extended Cox model is the best model compared to the Cox model.

Table 4.3. *Estimates of coefficient of time varying covariate (β_7).*

		Incorrect model			Correct model		
R	n	Bias	*SE*	MSE	Bias	*SE*	MSE
500	500	0.9211	0.1360	0.9311	0.0090	0.1546	0.1549
	1000	0.9133	0.0935	0.9181	0.0061	0.1081	0.1083
1000	500	0.9129	0.1360	0.923	0.0170	0.1537	0.1546
	1000	0.9106	0.0935	0.9154	0.0103	0.1075	0.108

A box plot has been also shown for the estimates of time varying covariate "plan" in both models for both 500 and 1000 iterations as shown in figure 4.1 and it shows the difference in the median of the estimate of the coefficient of covariate. The median of the incorrect model is greater than the median of the correct model.

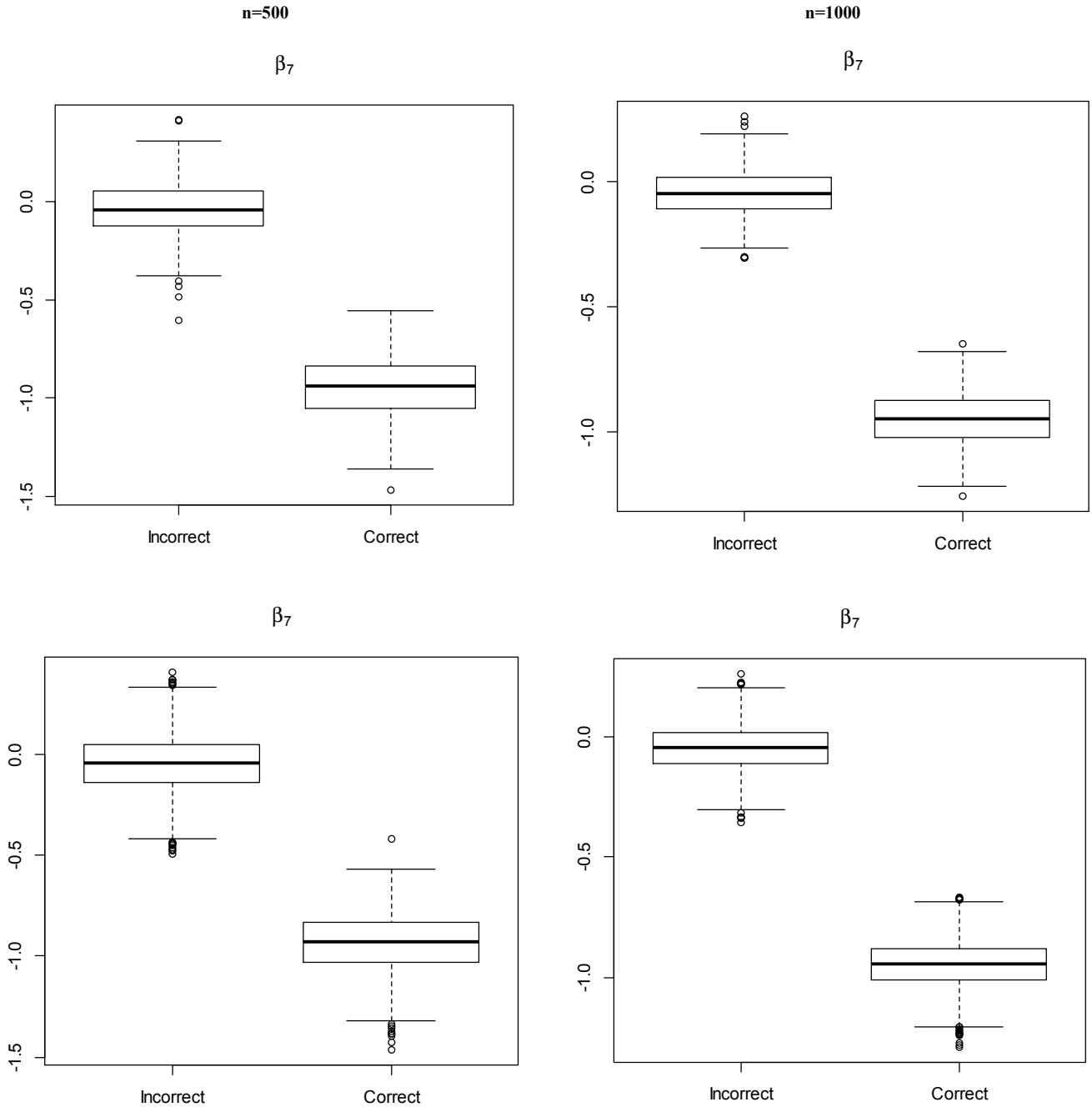

Figure 4.1. *A box plot of time varying covariate for incorrect and correct model.*

5. Conclusion and Recommendations

This study presented a method of modeling telecom customer retention using survival analysis technique. A set of seven covariates which consists of customer's sex, age, marital status ,employment status, customer's residence, number of subscriptions, customer's plan rate have been generated and "PermAlgo" package has been used to generate survival times conditional to this set of covariates. It was found those customer's demographic, socio-economic and behavioral characteristics affect the likelihood of churning. Two models (incorrect model and correct model) were fitted to determine which model analyzes well customer retention. The results highlighted the usefulness and the effect of the time dependent covariate in the Cox regression model; that is, by assuming the time dependent covariate (TDC) to be time independent covariate (TIC), the covariate became insignificant when it was significant and hence we can end up by making miscellaneous conclusions. A box plot has been used to show the difference in the two models for estimates of coefficients of time varying covariate (plan rate). It was found that the median of time varying covariate was smaller than the median of time independent covariate. We also

assessed the effect of the sample size on the model; as expected; the results indicated that the bias, the standard error, the mean squared error decreases as the sample size increases and the covariates become more significant as the sample size increases. In summary, the extended Cox model was the best description of how customer retention is achieved.

Most telecommunications companies in developing countries do not take into account the role of statistical data to achieve customer retention and may unwillingly make wrong decisions which can result into companies' losses. One limitation of this study was that companies do not keep proper customers' records; hence we recommend that consistent and well organized records which include all possible customers' demographic and socio-economic characteristics should be kept for every customer for better monitoring and evaluation in order to achieve customer retention. Given the usefulness of time effects in Cox model and the fact that more customers tend to change their rate plan or any other variable over time, we should analyze customer retention using extended Cox model instead of Cox regression model.

Appendix A: data set 1

Id	Event	Fup	Start	Stop	Sex	Age	Mar.stat	Resid	Employ	Subs	Plan
1	0	8	0	1	0	27	1	0	0	2	4
1	0	8	1	2	0	27	1	0	0	2	2
1	0	8	2	3	0	27	1	0	0	2	2
1	0	8	3	4	0	27	1	0	0	2	3
1	0	8	4	5	0	27	1	0	0	2	2
1	0	8	5	6	0	27	1	0	0	2	2
1	0	8	6	7	0	27	1	0	0	2	4
1	1	8	7	8	0	27	1	0	0	2	2
2	0	5	0	1	1	37	1	0	1	2	3
2	0	5	1	2	1	37	1	0	1	2	1
2	0	5	2	3	1	37	1	0	1	2	2
2	0	5	3	4	1	37	1	0	1	2	3
2	0	5	4	5	1	37	1	0	1	2	4
3	0	7	0	1	0	40	0	1	1	3	3
3	0	7	1	2	0	40	0	1	1	3	4
3	0	7	2	3	0	40	0	1	1	3	2
3	0	7	3	4	0	40	0	1	1	3	2
3	0	7	4	5	0	40	0	1	1	3	4
3	0	7	5	6	0	40	0	1	1	3	3
3	0	7	6	7	0	40	0	1	1	3	3
4	0	18	0	1	0	63	1	0	0	2	2
4	0	18	1	2	0	63	1	0	0	2	3
4	0	18	2	3	0	63	1	0	0	2	1
4	0	18	3	4	0	63	1	0	0	2	2
4	0	18	4	5	0	63	1	0	0	2	4
4	0	18	5	6	0	63	1	0	0	2	2
4	0	18	6	7	0	63	1	0	0	2	2
4	0	18	7	8	0	63	1	0	0	2	4
4	0	18	8	9	0	63	1	0	0	2	4
5	0	4	1	2	1	32	1	1	0	3	2
5	0	4	2	3	1	32	1	1	0	3	1
5	1	4	3	4	1	32	1	1	0	3	3
6	0	7	0	1	1	64	1	1	1	2	3
6	0	7	1	2	1	64	1	1	1	2	4
6	0	7	2	3	1	64	1	1	1	2	4
6	0	7	3	4	1	64	1	1	1	2	3
6	0	7	4	5	1	64	1	1	1	2	4
6	0	7	5	6	1	64	1	1	1	2	1
6	0	7	6	7	1	64	1	1	1	2	2
7	0	11	0	1	1	34	1	1	1	1	4

Appendix B: dataset 2

Id	Event	Fup	Start	Stop	Sex	Age	Mar.stat	Resid	Employ	Subs	Plan
1	0	8	0	1	0	27	1	0	0	2	4
1	0	8	1	2	0	27	1	0	0	2	4
1	0	8	2	3	0	27	1	0	0	2	4
1	0	8	3	4	0	27	1	0	0	2	4
1	0	8	4	5	0	27	1	0	0	2	4
1	0	8	5	6	0	27	1	0	0	2	4
1	0	8	6	7	0	27	1	0	0	2	4
1	1	8	7	8	0	27	1	0	0	2	4
2	0	5	0	1	1	37	1	0	1	2	3
2	0	5	1	2	1	37	1	0	1	2	3
2	0	5	2	3	1	37	1	0	1	2	3
2	0	5	3	4	1	37	1	0	1	2	3
2	0	5	4	5	1	37	1	0	1	2	3
3	0	7	0	1	0	40	0	1	1	3	3
3	0	7	1	2	0	40	0	1	1	3	3
3	0	7	2	3	0	40	0	1	1	3	3
3	0	7	3	4	0	40	0	1	1	3	3
3	0	7	4	5	0	40	0	1	1	3	3
3	0	7	5	6	0	40	0	1	1	3	3
3	0	7	6	7	0	40	0	1	1	3	3
4	0	18	0	1	0	63	1	0	0	2	2
4	0	18	1	2	0	63	1	0	0	2	2
4	0	18	2	3	0	63	1	0	0	2	2
4	0	18	3	4	0	63	1	0	0	2	2
4	0	18	4	5	0	63	1	0	0	2	2
4	0	18	5	6	0	63	1	0	0	2	2
4	0	18	6	7		63	1	0	0	2	2
4	0	18	7	8	0	63	1	0	0	2	2
4	0	18	8	9	0	63	1	0	0	2	2
4	0	18	9	10	0	63	1	0	0	2	2
4	0	18	10	11	0	63	1	0	0	2	2
4	0	18	11	12	0	63	1	0	0	2	2
4	0	18	12	13	0	63	1	0	0	2	2
4	0	18	13	14	0	63	1	0	0	2	2
4	0	18	14	15	0	63	1	0	0	2	2
4	0	18	15	16	0	63	1	0	0	2	2
4	0	18	16	17	0	63	1	0	0	2	2
4	0	18	17	18	0	63	1	0	0	2	2
5	0	4	0	1	1	32	1	1	0	3	3
5	0	4	1	2	1	32	1	1	0	3	3
5	0	4	2	3	1	32	1	1	0	3	3
5	1	4	3	4	1	32	1	1	0	3	3
6	0	7	0	1	1	64	1	1	1	2	3
6	0	7	1	2	1	64	1	1	1	2	3
6	0	7	2	3	1	64	1	1	1	2	3
6	0	7	3	4	1	64	1	1	1	2	3
6	0	7	4	5	1	64	1	1	1	2	3
6	0	7	5	6	1	64	1	1	1	2	3
6	0	7	6	7	1	64	1	1	1	2	3
7	0	11	0	1	1	34	1	1	1	1	4

References

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

[2] Ali, T. J. (2009), Predicting customer churn in the telecommunications service providers, Master's thesis, LuLea University of Technology.

[3] Collett, D. (2003), Modelling survival data in medical research, *CRC press*.

[4] Dudde, P. (2014), 'Rwanda-telecoms, mobile and broadband-market insights and statistics', *research and markets*.

[5] Fox, J. (2002), 'Cox proportional-hazards regression for survival data', *An R and S-PLUS companion to applied regression* pp. 1–18.

[6] Fox, J. & Weisberg, S. (2011), 'Cox proportional-hazards regression for survival data in r. an appendix to an r companion to applied regression'.

[7] Kleinbaum, D. G. (1998), 'Survival analysis, a self-learning text', *Biometrical Journal* 40(1), 107–108.

[8] 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* pp. 114–27.

[9] Martinussen, T. & Scheike, T. H. (2007), *Dynamic regression models for survival data*, Springer Science & Business Media.

[10] Ocloo, C. E. & Tsetse, E. K. (2013), 'Customer retention in the ghanaian mobile telecommunication industry', *European Journal of Business and Social Sciences* 2(7), 136–160.

[11] Portela, S. & Menezes, R. (2011), 'detecting customer defections: an application of continuous duration models', *Journal of Global Strategic Management* (09), 22–30.

[12] RDB (2013), "The national customer satisfaction survey", Rwanda Development Board.

[13] Reichheld, F. F. & Schefter, P. (2000), 'E-loyalty', *Harvard business review* 78(4), 105–113.

[14] RURA (2015), 'Communication and media', *Rwanda Utilities Regulatory Authority* .

[15] Sharmeela-Banu, S. A., Gengeswari, K. & Padmashantini, P. (2012), 'Customer retention practices among the major retailers in malaysia', *International Journal of Academic Research in Business and Social Sciences* 2(6), 157–166.

[16] Timothy, K. L., Cooil, B., Aksoy, L., Andreassen, T. W. & Weiner, J. (2007), 'The value of different customer satisfaction and loyalty metrics in predicting customer retention, recommendation, and shareof-wallet', *Managing Service Quality: An International Journal* 17(4), 361–384.

[17] Van Den Poel, D. & Lariviere, B. (2004), 'Customer attrition analysis for financial services using proportional hazard models', *European journal of operational research* 157(1), 196–217.

[18] Williams, C., Solomon, G. & Pepper, R. (2012), 'The impact iof mobile telephony on economic growth. a report for groupe speciale mobile association (gsma)'. A report for GSM Assocition.

[19] Wong, K. K.-K. (2011), 'Using Cox regression to model customer time to churn in the wireless telecommunications industry', *Journal of Targeting, Measurement and Analysis for Marketing* 19(1), 37–43.