Predicting Insurance Policyholders Attrition In Nigeria Insurance Firm:-A Cox – Regression Approach (Implemented Using R-Codes)

A.B. Sogunro¹, Hamadudallah², R.O. Ayorinde³

Department of Actuarial Science & Insurance, University of Lagos, Nigeria.

Corresponding author: A.B. Sogunro

ABSTRACT: Insurance is faced with hostile relationship between the insurer and the insured more than most business due to mere redistribution of money rather than providing a factual product or service to the insured. These lopsided attitudes make the insurer appear feeble in the presence of policyholders and they sought value elsewhere that can meet or promise to meet their current or future needs. In the light of this problem faced by policyholders, there is a need to construct a viable churn predictive model with proper ability of identifying any policyholder with high-risk of cancelling policy. This study used record data of one Nigerian insurance company for the construction of the model with a Cox-regression approach, implemented using R-Codes which is capable of investigating concurrently the effect of several explanatory variables on the cancellation of policy by policyholders. The result of the analysis shows that marital status and sum assured have potential influences on churn out of the several key candidate predictors. Hence, insurance marketing teams should look onward at how they can most effectively put up physically powerful relationships with their customers and preserve against declining retention.

KEYWORDS-Attrition, BaselineChurn, hazard, Cox regression, Modeling, Proportional Hazard.

Date of Submission: 20-11-2018 Date of acceptance: 06-12-2018

I. INTRODUCTION

1. Background of the study

Plummeting customer attrition and the ability to foresee that a particular customer is at high risk of churning while there is still time to do something about it are key business goals of every business organization. While an organization understanding of customers’ needs, preferences, sentiments, behaviour, propensity and engagement to change which has a significant impact on customer lifetime value and on the company’s objective has become one of the top issues in business environment.

Financial sectors globally, including insurance has become highly competitive due to the liberalization of financial systems to speed up the process of economic growth [1]. Insurance improves the investment environment and contributes to the economic growth of businesses and economies in general by protecting the financial health of businesses and encouraging home production and trade [2]. While Life insurance on the other hand, a branch of insurance, ensures financial security to family (spouse, children and dependent parents), help to prepare for life’s uncertainties and give peace of mind to the dependants. People depend on insurance for many reasons, but the basic principle is to minimizing personal risk on a day-to-day basis. Insurer has the commitment to provide the corresponding economic compensation in return to the premium payment made by the insured in advance for cover against risk that a particular event would occur [3]. The importance of insurance sector is not only derived from provision of cover against risks that citizens, organizations have to face, but also from their specific business activity.

Insurance companies have over the years expectedly devoted attention to the development of products for individuals and corporate firms, and struggled to acquire new customers for the developed products but showed less concern about the old clients. As the insurers cover risks, wrestle with liquidity and then tries to survive in changing market environment, general economic regression and deregulation, challenges in a market characterised by product saturation, soft pricing, competitive premiums, rise of connected technologies and online comparison tools and competition with each other in order to grow their business, it is imperative that the company do not let customer service suffer. They should not fail to engage their customers on an individual basis, and realize those satisfied policyholders are a company backbone. The insurers should ensure that sophisticated customer relationship management and churn management techniques which can attract new customers, reflex in engaging policyholders from other competitors and prevent them to leave or cancel their policy are employed. Even though insurance is complex, customers want to be involved, emotionally and rationally in order to raise their share.
Predicting Insurance Policyholders Attrition In Nigeria Insurance Firm: A Cox – Regression …

Studies review that customer defection is due to relative ease and low cost of changing insurer, offering lower costs price and lacking basic customer value, spending a lot of money and effort on advertising services, and honouring existing no-claims bonuses offer to the new policyholders when it comes to picking an insurance policy that satisfied their requirement in order to increase incentives to consumers to change, while frequently leaving existing customers on the receiving end of poor customer service of which existing customers are most valuable assets[4]. They reciprocate people for being unfaithful and leaving opponent company, rather than cultivating faithful customers which will enhance revenue.

Customer attrition which can be voluntary (personal decision by the customer to switch to another company or service provider), or involuntary (a switch due to circumstances such as death or relocation to a distant location), is not just about plummeting costs to perk up revenue levels. Companies must know that there is a separate requirement to re-balance funds spent on acquiring new customers in resistance to retaining the faithful ones they have already. Much effort most not only be put on persuading clientele to sign contract by business organizations, but also ensure that the existing clients are retained. Some insurance industry lack proper management due to lack of transparency which had led to customers losing their money in the process and thus making the public lose trust in the industry [5]. Customer attrition can be used as an indicator of customer satisfaction[6], and help organizations to detect early warning signs such as reduced transactions or negative experiences. Organisations often use customer attrition analysis and customer attrition rates as one of their key business metrics along with cash flow earnings before interest and tax e.t.c. because an organisation spend more than five times as much to obtain a new customer than to retain an existing ones, and it costs far more to re-acquire deflected customers [7]. Hence, companies go to great lengths to keep customers on board.

Therefore, to be successful at retaining policyholders who would otherwise ditch the policy, insurance vendor needs to know the buyers’ satisfaction level for an insurance product accurately, and adopt a new way of harnessing their volumes of customer and business information in order to move ahead by analyzing historical record of what has happened and create models that can predict what will happen next in future. This will in the long run reduce or eliminate a large proportion of policyholder to cancel policy. This study is therefore carried out to construct a viable churn predictive model with proper ability of identifying any policyholder with high-risk of cancelling policy. The study employed a statistical approach to estimate the effect of behavioural and demographic characteristics of policyholders on their aim to cancel their policy with an insurance firm implemented using R- Codes which possesses extensive and powerful abilities that are tightly linked with its analytic abilities.

II. PREVIOUS WORKS

The growing awareness of different insurance companies in the resolution of preference of clients for a variety of insurance plans has opened up new areas of study in Nigeria insurance market. There has been limited research on policyholder’s attrition for insurance companies in Nigeria. Numerous pragmatic studies and models have confirmed that customer attrition remains one of the major incinerasors of organisation value, and has been modelled in the telecommunication industries, financial institutions, retail markets e.t.c. using either discrete or continuous time by applying different Data mining techniques such as logistic regression, Cox regression, Neural Network Ordinal Regression, Decision trees, Bayesian networks, AdaBoosting, Proportional hazard model, Stratified Cox – Proportional Hazards model etc. using demographic and behavioral attributes of the customers.

[8], employed data mining methods for Customer churn management (CCM) by first selecting customers with equal characteristics using clustering K-means method and secondly make use of churn index and decision tree CART to analyze and know reasons of customer churn. The data mining process was done by Clementine software on set of data gathered from seven “Iran Insurance” branches in Anzali as population size while, Customer clustering and knowing the reasons of churning by decision tree CART in order to make company choose better policy to reduce that.

Also, [9], exploit logistic longitudinal regression model that incorporates time-dynamic explanatory variables to predict individual customers’ risk of leaving an insurance company by fitting interactions to the data. As an intermediate step in the modelling procedure, they apply generalised additive models to identify non-linear relationships between the logit and the explanatory variables. Their result shows that the model performs well in terms of identifying customers that are likely to leave the company each month using both out-of-sample and out-of-time prediction indication, and their approach is general and may be applied to other industries as well.

In addition, [10], employed in their study the better applicability and efficiency of hierarchical data mining techniques in predicting those customers at high risk of leaving the company. They considered three hierarchical models (ANN + ANN + Cox, SOM + ANN + Cox, and α - FCM + ANN + Cox) by combining four different data mining techniques (backpropagation artificial neural networks (ANN), self-organizing maps
(SOM), alpha-cut fuzzy c-means (α-FCM), and Cox-proportional hazards regression model) for churn prediction. The first part of the models aims to cluster data in two churner and non-churner groups and also shift out misrepresentative data or outliers, while the clustered data as the outputs are used to allocate customers to churner and non-churner groups by the second technique. The correctly classified data are used to create Cox proportional hazards model which performance evaluation was considered using an Iranian mobile dataset. Their experimental results show that the hierarchical models outperform the single Cox-regression baseline model in terms of prediction accuracy. Types I and II errors, RMSE, and MAD metrics. They concluded that, the α-FCM + ANN + Cox model significantly performs better than the two other hierarchical models.

Other data mining techniques successfully applied in customer attrition prediction include:[11], [12], [13], [14],[15],[16],[17], [18], [19],[20], [21],[22], [23], [24], [25] and [26].

Also, ordinal regression technique was exploit by [27].Cox regression [28], Neural network and Cox regression [29], Stratified Cox Proportional Hazards model [30], Proportional hazard model [31], Artificial neural networks [32] and [33], [34].While[35], [36],[37], [38],[40], [41], and [42] applied logistic regression duration models but, their approach cannot assess association between survival time and covariates.

More also, ADTreesLogit technique was used by [43], AdaBoosting [44], Bayesian networks [45] and[46], and Decision trees [47] to forecast customer attrition.

III. RESEARCH METHOD

3.1 Data Selection

The potential influences on attrition of several key candidate predictors: age; sex; marital status and sum assured (demographic and behavioural attributes) of 5,958 life insurance policyholders’ records were extracted directly from data information of one of the well-known Nigeria insurance company to carry out the research analysis for the study.

The policyholders were selected for observation over a period of 36,004 days starting from the policies activation year April 01, 2000 till the end of the observation period in December 31, 2012. The Customer demographic and behavioral attribute data which can be classified as static and temporal respectively are divided into dependent and independent variables. The dependent (criterion variable) called the status variable identifies whether the event (attrition) has occurred for a given case if not, it said to be censored. Censored cases are not used in the computation of the regression coefficients, but are used to compute the baseline hazard.

The status of policyholders represent dependent variable, while independent variables are further grouped into two; the commonly used demographic variable (age, sex and marital status) and sum assured (behavioral attribute) of the policyholders. These two groups of the independent variables play an important role in influencing policyholder’s intention to cancel the policy, hence modeling analysis include these factors.

To know the effect of these variables on policyholder cancellation rate, Cox-regression which attempt to model a generating function which underlines data and capable of investigating concurrently the effect of several explanatory variables on the cancellation of policy by policyholders was employed and implemented using R-Codes.

3.2 Analysis and Results

In analysing the effect of the factors on cancellation of the policy, random samples of policyholders are selected with their time spent as policyholder including active members and various demographic attributes. The observation period which is the difference between the issue date of the policy and the cancellation date of the policy are converted to days. The covariates (Sex and Marital Status) are coded by attaching a numerical values to each as follows: Sex ( Male = 1 and Female = 2), Marital status (Single =1 and Married=2), while sum assured are grouped into{ (0 – 299,999) = 1, (3,000,000 – 999,999) = 2, (6,000,000 – 999,999) = 3, (10,000,000 – 99,999,999) = 4 and (100,000,000 – 500,000,000) = 5 } and age in numeric value. The dependent variables which represent the survival time of the policyholders are also grouped into three and then coded as follows {active = 0, maturity = 1, while surrender and death = 2}. Cox regression analysis was then run using status as dependent variable with code of interest on it.

1.3. Estimation Procedure

The minimum requirement for Akaike Information Criteria (AIC) value needed by the model to perform effectively was generated by the system during the estimation procedure in order to eliminate variable having AIC lower than the minimum required AIC value. Sex and age with 5074.3 and 5073.5 AIC values respectively are eliminated during the first and second estimation procedure process. This implies that the two variables do not have any contribution (insignificant to the dependent variable.

In the last estimation process, new AIC value 5073.5 was generated and both marital status and sum assured had AIC value greater than the estimated AIC value making the two variables not to be eliminated. This indicate that marital status and sum assured both with 5079.5 and 5120.0 AIC value respectively have highly
contributed in significance to the proposed model. The median value at time 2306 days is 0.05961, and coef, exp(coef), se(coef), z and p values for both marital status and sum assured are [coef = 0.482, exp(coef) = 1.62, se(coef) = 0.1774, z = 2.72, p = 6.5 \times 10^{-13}] and [coef = -0.386, exp(coef) = 0.68, se(coef) = 0.0599, z = -6.45, p = 1.1 \times 10^{-10}] respectively. The result of the estimation procedure is given below.

<table>
<thead>
<tr>
<th>Hazardtime</th>
<th>Min.</th>
<th>Min.</th>
<th>1st Qu.</th>
<th>1st Qu.</th>
<th>Median</th>
<th>Median</th>
<th>Mean</th>
<th>Mean</th>
<th>3rd Qu.</th>
<th>3rd Qu.</th>
<th>Max.</th>
<th>Max.</th>
<th>Start: AIC=5076.25</th>
<th>Surv(tday, status != 0) ~ sex + age + marital + sumas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.:0.00000</td>
<td>Min. : 30</td>
<td>1st Qu.:0.02232</td>
<td>1st Qu.: 1588</td>
<td>Median:0.04323</td>
<td>Median : 2306</td>
<td>Mean :0.05961</td>
<td>Mean : 2817</td>
<td>3rd Qu.: 0.10719</td>
<td>3rd Qu.: 3385</td>
<td>Max. : 0.12501</td>
<td>Max. : 11717</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[Surv(tday, \text{status }! = 0) \sim sex + age + marital + sumas\] (1)

\[DfAIC\]
- sex 1 5074.3  
- age 1 5075.5  
<none> 5076.2  
- marital 5082.0  
- sumas 5122.3

\[Step: \text{AIC } = 5074.31\]

\[Surv(tday, \text{status }! = 0) \sim age + marital + sumas\] (2)

\[DfAIC\]
- age 1 5073.5  
<none> 5074.3  
- marital 5080.1  
- sumas 5120.3

\[Step: \text{AIC } = 5073.5\]

\[Surv(tday, \text{status }! = 0) \sim marital + sumas\] (3)

\[DfAIC\]
<none> 5073.5  
- marital 5079.9  
- sumas 5120.0

\[Call:\]
\[\text{coxph(formula } = \text{Surv(tday, status }! = 0) \sim marital + sumas, data = X)\]

\[\text{coef}\ exp(\text{coef}) \ se(\text{coef}) \ z \ p\]
marital0.482 1.62 0.1774 2.72 6.5e-03
sumas -0.386 0.68 0.0599 -6.45 1.1e-10

\[\text{Likelihood ratio test } = 55.8 \text{ on 2 df }, p = 7.8e - 13, \quad n = 5958, \text{number of events } = 316\]

IV. DISCUSSION OF THE RESULT

During the study period of 5,958 policyholders, 316 have surrendered their policies before the end of the follow-up study due to one reason or the others. The results of the analysis indicated that marital status and sum assured have direct influence on attrition of policies. This implies that marital status and sum assured have significant contribution to policy cancellation and influence the risk of the policyholders to leave or cancel their policy. The Table1 below summarised the output result from the model-generation process with baseline of 0.0596.
Predicting Insurance Policyholders Attrition In Nigeria Insurance Firm: A Cox – Regression ...

Table 1 Summary of the Analysis

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Coefficient</th>
<th>Exponential coefficient</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marital status</td>
<td>0.482 (β)</td>
<td>1.62</td>
<td>0.0065</td>
</tr>
<tr>
<td>Sum assured</td>
<td>-0.386 (α)</td>
<td>0.68</td>
<td>1.1 x 10^-11</td>
</tr>
</tbody>
</table>

Likelihood ratio test = 58.8 on 2 df, p value 7.8 * 10^-13, n = 5956, event number = 316
Mean hazard = 0.05961 at time 2917, Median hazard = 0.04323 at time 2306

Hence, the model in equation 5, represents the formulated generating function which underlies the data of the policyholders with a Cox-regression approach, implemented using R-Codes.

\[
h(t, \text{marital status}, \text{sum assured}) = h_0(t) e^{(\beta \cdot \text{marital status} + \alpha \cdot \text{sum assured})} = 0.05961 e^{(0.482 \cdot \text{marital status} - 0.386 \cdot \text{sum assured})}
\]

(5)

The survival curve below (fig. 1) depicts the proportion of the population of policyholders still on the policy after a given time. It represents the experience of each policyholders and probability of survival from the time the observation started. Thus, any point on the survival curve gives the probability that the "average" policyholder will remain a customer past that time.

![Figure 1 Graph of the Survival Function](image)

The survival graph shows that between the beginning and 332 days of the follow-up study periods, there are 94.6% chances (95% confidence interval) that the policyholders will not churn their policy. The rate of churning increases between the period 332 and 478 days and then reduces thereafter with 66% chances of survival till 541 days. The rest of the policyholders were maintained with 57% chances of survival till end of the follow-up study.

The basic hazard curve depicted below in Fig. 2, illustrates the probability of failure at time \( t \) given survival until time \( t \). The curves shows a visual representation of the effect of policyholders and predicts potential to churn for the "average" policyholders.

![Figure 2 Graph of the Hazard Function](image)

Past 350 days the hazard curve in fig. 2 above, becomes less smooth which implies that there are fewer customers who have been with the insurance firm for that long.
V. CONCLUSION AND RECOMMENDATION

Regardless of the line of business, a more effective approach to improve customer retention is to predict potential policyholder cancellation by building a model with high quality score that can reasonably catch and explain the factors which influence a policyholder to cancel the policy or leave the company. This makes it challenging for the insurer to get a holistic understanding of their policyholders, detect early warning signs and engage them with retention offers. Therefore, by deploying the technologies (Attrition prediction models) coupled with effective retention programs, customer attrition could be better managed to stem the significant revenue loss from defecting policyholders [48].

Hence, this study recommends the need for Nigeria insurance companies to conduct continuous customer service satisfaction surveys on their product, and also put emphasis on the aspects of the service or product that seemed rather dissatisfying with a view to making an improvement on them. Also, an international standards regarding recognition and sharing of global best practice in customer service can be developed in order to reduce customer attrition. This will strategically align organization to focus on delivering excellence in policyholder service as well provide recognition of success through a third party registration scheme.

More also, insurers should focus on customer care challenges (e.g. change in capability, delivery of consistent customer experience, talent churn, inefficient processes and technology obsession ways), looking at range of ways to change the operating model to ensure satisfied customers. They should also budge from product-oriented marketing strategy to a more customer-oriented one to be able to uphold client loyalty, reduce churn and grow their objective naturally.

In addition, satisfaction survey regarding insurance service delivery to policyholders should always be carried out to know the mind set of policyholder on the services rendered to them by the insurer and also identify area insurers should improve on for serving policyholder.

In conclusion, companies can make a diagnosis and mend the starting place of customer departures and dissatisfaction by radically re-thinking the rationale for their customer service expense – it’s not about reducing customer service costs, it’s about stopping churn and thereby having a positive impact on top-line revenue (Genesys, 2013). The business organization must also focus and force all the member of staff from the most senior to the least junior of the management hierarchy to take consumer service and satisfaction as their priority in all the frame work of policies, practices and information. It requires continual monitoring and experience examination, opinions and potential customers.

REFERENCES

Predicting Insurance Policyholders Attrition In Nigeria Insurance Firm:-A Cox – Regression Approach

A.B. Sogunro


