Journal of the Operational Research Society

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjor20

# Estimating customer churn under competing risks

Pallav Routh , Arkajyoti Roy & Jeff Meyer

**To cite this article:** Pallav Routh , Arkajyoti Roy & Jeff Meyer (2020): Estimating customer churn under competing risks, Journal of the Operational Research Society, DOI: 10.1080/01605682.2020.1776166

**To link to this article:** https://doi.org/10.1080/01605682.2020.1776166

JOURNAL OF THE OPERATIONAL RESEARCH SOCIETY
https://doi.org/10.1080/01605682.2020.1776166

ORIGINAL ARTICLE

# Estimating customer churn under competing risks

Pallav Routh[a], Arkajyoti Roy[a] and Jeff Meyer[b]

[a]The University of Texas at San Antonio, TX, USA; [b]Bowling Green State University, Bowling Green, OH, USA

**ABSTRACT**
Customer churn management focuses on identifying potential churners and implementing incentives that can cure churn. The success of a churn management program depends on accurately identifying potential churners and understanding what conditions contribute to churn. However, in the presence of uncertainties in the process of churn, such as competing risks and unpredictable customer behaviour, the accuracy of the prediction models can be limited. To overcome this, we employ a competing risk methodology within a random survival forest framework that accurately computes the risks of churn and identifies relationships between the risks and customer behaviour. In contrast to existing methods, the proposed model does not rely on a specific functional form to model the relationships between risk and behaviour, and does not have underlying distributional assumptions, both of which are limitations faced in practice. The performance of the method is evaluated using data from a membership-based firm in the hospitality industry, where customers face two competing churning events. The proposed model improves prediction accuracy by up to 20%, compared to conventional models. The findings from this work can allow marketers to identify and understand churners, and develop strategies on how to design and implement incentives.

**ARTICLE HISTORY**
Received 29 January 2019
Accepted 25 May 2020

**KEYWORDS**
Competing risk; random survival forest; customer churn management; machine learning

## 1. Introduction

Managing customer churn is essential for developing and maintaining customer relationships (Ascarza et al., 2016). The first step to cure customer churn is to predict which customers have a high probability of ending their relationship with the firm (Neslin et al., 2006). The second step is to target the predicted at-risk customers using incentives such as pricing offers (Ascarza et al., 2016) or communications such as emails (Ascarza, Netzer, & Hardie, 2018). Models that accurately predict customer churn are pivotal in targeting the right customers, thereby decreasing the cost of the marketing campaign and using scarce firm resources more efficiently (Neslin et al., 2006).

Traditional models that predict customer churn assume (i) there is a single event of churn, (ii) the model's functional form is known and correctly specified, (iii) some underlying distribution to model risk based on customer behaviour, or (iv) a lack of temporal information on the event of churn. However, in many firms there exists multiple outcomes of churn (Braun & Schweidel, 2011). Furthermore, in various real-world applications, the functional form is unknown or incorrectly specified, and the underlying distributional assumptions are violated (Bodapati & Gupta, 2004). Moreover, the changes in risk of churn across time is very important for a marketer, especially when also observing customer behaviour. For example, the behaviour of a customer who churned after a week means less to the decision-maker compared to the behaviour of a customer who churned after a year. These limitations become even more acute in econometric models when there are a large number of parameters (Wedel & Kannan, 2016). The traditional churn prediction models provide a research gap for more robust methods that can help marketers better understand the relationships between churn and behaviour, and implement more accurate marketing campaigns.

Existing studies have analysed various aspects of customer behaviour that affect customer churn. Logistic regression models have been used to predict events of churn (Neslin et al., 2006). For multiple events, multinomial logistic regression can be used instead of binary logistic regression. However, there are limitations to logistic regression in the context of churn prediction, including ignoring time to churn (Menard, 2002). Alternatively to account for time to churn, survival models have been developed. For example, Reinartz et al. (2005) used a customer's frequency of purchase and cross-buying tendencies to model customer churn. When modelling a customer's probability to churn, these customer-specific behaviours (i.e. the model's covariates) can define the behavioural, demographic, geographic, or

CONTACT Arkajyoti Roy arkajyoti.roy@utsa.edu The University of Texas at San Antonio, TX, USA

psychographic composition of a customer. The probability predictions from a customer churn model can then be used to identify customers for churn management (Lemmens & Croux, 2006). Specifically, a marketer can identify the covariates that impact the probability to churn, and then implement incentives designed to affect those covariates, directly influencing customer behaviour. For instance, shipt.com, a grocery delivery service provides free delivery for orders exceeding $35, and various retail services provide incentives such as follow-up discounts that expire before a specific date. For shipt.com, the incentive is designed to increase customer expenditure, whereas for the retailers, the incentive is designed to decrease time between purchases. Both incentives affect customer behaviours that influence customer churn, but these marketing decisions are only possible after accurately establishing the relationships between the risk of churn and customer behaviour. In the presence of multiple events of churn and a large number of covariates, identifying the correct relationships poses a significant challenge for marketers.

Given these challenges, the two fundamental research questions we study in this work are *(a) how to accurately and robustly predict customer churn for competing outcomes, and (b) how do different aspects of customer behaviour affect different outcomes of churn?* We address these questions through our primary contribution of a churn prediction methodology, which

i. accounts for multiple events of churn using a competing risk framework,
ii. employs a random survival forest (RSF) that can model time to churn,
iii. is robust to distributional assumptions and a lack of functional form for modelling, and
iv. provides tools that can explore relationships between behaviour and churn.

This methodology can aid decision-making by improving accuracy of churn predictions and relationships between behaviour and churn.

The remaining portion of this paper is structured as follows. In the next section, we review existing literature at the intersection of churn management, competing risk, and machine learning for churn prediction. In Section 3, we provide details of the data, including covariates used in modelling. In Section 4, we develop the competing risk random survival forest model for customer churn predictions. In Section 5, we present an exploratory analysis of the results and evaluate the performance of the proposed methodology. In Section 6, we discuss the implications of this research. Finally, we discuss the limitations of this work followed by concluding remarks.

## 2. Literature review

A common strategy for firms to reduce churn is to provide incentives to customers likely to churn so the customers maintain their relationship with the firm (Neslin et al., 2006). It is also common for such firms to design the incentives so that specific aspects of customer behaviour is affected (Ascarza, Neslin, Netzer, et al., 2018). For instance, a buy-one-get-one-free incentive may be designed to promote cross-buying across different categories of brand, while a flash sale may be designed to shorten the time between purchases. We posit that if different aspects of customer behaviour affects different avenues of churn, then designing such campaigns should differ based on the churning event. This approach is especially relevant for firms where customers can opt for different avenues of churn. Our work evaluates this hypothesis through a modelling framework that builds upon three areas of existing research: customer churn management programs for multiple events, estimation of risk of churn under competing events, and prediction models estimating customer churn.

### *2.1. Customer churn management*

Customer churn refers to the decision of a customer to end her relationship with the firm (Lemmens & Croux, 2006). Churn is the antithesis of retention, and it often serves as the basis for the CLV framework (Gupta et al., 2004). However, the importance of measuring churn and retention goes beyond CLV. Retention also drives firm profitability (Gupta et al., 2004) and serves as an indicator of firm strength (Ascarza, Netzer, & Hardie, 2018). Therefore, customer churn management is critical for firm success.

Firms engage in both proactive and reactive customer churn management (Ascarza, Neslin, Netzer, et al., 2018). In proactive management, firms embark upon actions to prevent the causes of churn. In reactive management, firms wait for churn to occur and then attempt to recapture the customer. We focus on proactive management, which requires the prediction of churn through the measurement and modelling of churn predictors (Neslin et al., 2006). In marketing literature, researchers have examined many predictors of customer churn across a variety of factors. For instance, researchers have examined the influence of marketing efforts, such as loyalty programs and promotions (Lewis, 2004). Another factor receiving attention is in the area of customer satisfaction and commitment (Gustafsson et al., 2005). Customer usage behaviour, such as purchase frequency and cross-buying (Reinartz et al., 2005), has also received much attention. Other factors include switching costs, customer characteristics, and social connectivity (Ascarza, Netzer, & Hardie, 2018).

These studies highlight that an important aspect of churn management programs is to predict which customers are most likely to churn by examining certain factors (or covariates) that contribute to churn. Usually, a marketer would employ a decision model that predicts the binary outcome of churn as a function of such factors. Based on the prediction, the marketer can then decide whether or not to provide incentives. This decision is crucial if firms want to retain the customer and drive profits from them in the future. However, consider a scenario when there are multiple outcomes of churn. The marketer's decision model would now differ from the previous scenario, since it needs to account for multiple or competing events of churn within one prediction model. Additionally, if the firm launches event specific churn incentives, the marketer's decision would be dependent on the event specific predictions from the decision model, because the firm would ideally want to provide incentives based on the event that is most likely to occur. Extant research shows the returns of event specific retention management strategies depend on variations in tenure and customer characteristics that are specific to a competing event (Braun & Schweidel, 2011). Therefore, under such circumstances, making the correct decisions would depend on a churn prediction framework that would allow for competing churning events.

### 2.2. Competing risk

Competing risk events arise when there are more than one avenue for a customer to end her relationship with the firm. That is, the customer can choose from one of many exit choices to churn away at any point in time during her relationship with the firm. A competing risk model computes the event-free churn probability of an individual, followed by considering either the event of interest or the competing event, thereby capturing the effects of the competing events. Based on the methodology introduced by Hoel (1972) using latent lifetimes to model the risks from competing events, Prentice and Gloeckler (1978) introduced the cause-specific hazard function (CSHF). This method can be used to model the effect of covariates on the CSHF estimates by extending the proportional hazard framework of Cox (1992). However, this method fails to incorporate the other competing events while estimating the risk from a specific event. To overcome this limitation, Fine and Gray (1999) developed the cumulative incidence function, which accounts for the effects from competing events and establishes relationships between the covariates and the cumulative incidence estimates of churn.

Traditionally, marketing researchers have used empirical frameworks such as logit methods that model a single event of churn. Such models can be extended to scenarios that involve decision making for multiple event or causes of churn. More recently Braun and Schweidel (2011) and Kumar et al. (2018) have employed a competing risk methodology for estimation of customer worth. Although such competing risk models can account for multiple events of churn, they rely on parametric data and known functional forms to model the relationships. If the distributional assumptions or functional forms are invalid, it could lead to inaccurate predictions, and therefore, long term losses for the firm. Thus, a churn prediction model that not only accounts for multiple events of churn, but also does not rely on a known or correctly specified functional form or underlying distributions is highly desirable.

### 2.3. Prediction models in customer churn

Customer churn has been modelled using logit models and machine learning methodologies, such as decision trees, especially when the event is binary and when modelling the duration is ignored (Lemmens & Croux, 2006). However, this class of models cannot account for time to churn and also suffers from high variance (Dudoit et al., 2002). Incorporating time to churn provides additional information for the prediction model. When modelling time of customer engagement is important, semi-parametric regression-based models have been used to model duration and churn (Bolton, 1998; Gönül et al., 2000; Knott et al., 2002; Levinthal & Fichman, 1988). Due to their regression-based modelling, these methods are prone to model mis-specification, especially under a large number of parameters (Bodapati & Gupta, 2004).

In order to avoid such mis-specifications, we rely on machine learning models that are non-parametric and makes no distributional assumptions. Tree based models such as decision trees are popular for classification tasks such as churn prediction. However, simple decision trees do not take into account the temporal aspect of churn, that is, the time to churn. A separate class of decision trees, called decision survival trees, was developed to handle survival data (LeBlanc & Crowley, 1995).

Table 1. Customer-specific covariate descriptions.

| Covariate | Description |
|---|---|
| **Transaction intensity covariates** | |
| expenditure | Total dollar amount spent by the customer |
| frequency | Frequency of purchases |
| avg_expenditure | Average expenditure per transaction in dollars |
| **Transaction timing covariate** | |
| avg_ipt | Average time per consecutive transactions in days |
| **Cross-buying covariates** | |
| areas_dining | If the customer has transacted in the dining areas |
| areas_golf | If the customer has transacted in the golfing areas |
| areas_marina | If the customer has transacted in the marina |
| **Donation covariates** | |
| christmas_donation | If the customer has donated to the firm during Christmas |
| scholarship_donation | If the customer has donated to the firm's scholarship fund |
| **Demographic covariates** | |
| yom | Years of membership in years |
| dist_km | Distance of the customer from the firm in kilometres |

Although survival trees can be used for churn prediction, they are still prone to high variance in their predictions which could reduce the accuracy of results.

Bagging and forests overcome the limitations of decision trees by minimising the variance from a single tree (Breiman, 1996, 2001). Bagged trees and forests are grown by drawing several bootstrap samples and then growing a single tree in each sample. Forests add a layer of randomisation by selecting random inputs for growing a tree. Therefore, a forest-based survival model that accounts for multiple or competing events of churn would be ideal for marketers that wish to make accurate predictions for event specific churn management campaigns.

Ishwaran et al. (2008) extended the forest algorithm to censored observations and referred to the approach as random survival forest (RSF). Furthermore, the survival ensemble was based on the non-parametric estimation of survival and hazard. More recently, Ishwaran et al. (2014) extended the application of RSF under a competing risk setting in health-care. In this study, we apply the RSF methodology developed by Ishwaran et al. (2014) to estimate the churn and retention probabilities of customers who face multiple competing events of churn from a membership-based firm.

## 3. Data description

The data used in this study come from a firm in the hospitality industry. The services provided by the firm are similar to private clubs, country clubs, and resorts. For this firm, membership is open to the general public. At any point in time, if a customer wishes to terminate his membership, he has two choices – he may choose to "*resign*" from the firm (i.e. terminate the membership permanently) or he may choose to take a "*leave of absence*" (i.e. terminate the membership temporarily). For brevity, henceforth, we refer to leave of absence as LOA. If a customer opts for LOA, he may return as an active customer at a reduced cost. On the other hand, if a customer resigns, he returns as a new customer. We model resign as the competing event and LOA as the event of interest in the competing risk analysis. Membership between the firm and its customer is contractual. Thus, the firm is notified immediately when a customer resigns or takes a leave of absence.

In addition to customer churn data, we also collected customer transaction information, which were recorded after a transaction was made. From this transactions database, we derived several customer behavioural measures, which we used as independent covariates in our model. We classify these covariates into four groups for ease of description: transaction intensity, transaction timing, cross-buying, and donation. In addition, due to the geographic location of the firm near the Upper Midwest and its proximity to a large lake, we established seasonal measures of the covariates to extract seasonal impact. Specifically, we construct off-season and peak-season variations of the transaction intensity and transaction timing covariates. The peak-season was identified as May to September by observing trends in data and validated through discussions with decision-makers at the firm. Table 1 shows the covariates used for modelling and their description. Here, we prefixed the words "seasonal_" and "non_seasonal_" to the covariates to denote their peak-season and off-season variants, respectively.

Transaction intensity covariates measure the magnitude of a specific transaction (Reinartz & Kumar, 2003; Reinartz et al., 2005; Venkatesan & Kumar, 2004). Total frequency (referred to as *frequency* in the variable description Table 1) is a transaction intensity covariate that computes the total number of purchases made by a customer within the firm over the entire duration of the study. Similarly, total expenditure (*expenditure*) represents the total dollar amount spent by the

**Table 2.** Covariate descriptive statistics.

| | Covariate | Mean | Standard.deviation |
|---|---|---|---|
| 1 | seasonal_avg_ipt | 28.70 | 50.26 |
| 2 | seasonal_avg_expenditure | 167.98 | 408.60 |
| 3 | seasonal_expenditure | 5539.69 | 9609.75 |
| 4 | seasonal_frequency | 52.03 | 57.57 |
| 5 | seasonal_areas_dining | 0.84 | 0.37 |
| 6 | seasonal_areas_marina | 0.34 | 0.47 |
| 7 | seasonal_areas_golf | 0.53 | 0.50 |
| 8 | non_seasonal_avg_ipt | 54.33 | 73.38 |
| 9 | non_seasonal_avg_expenditure | 344.23 | 732.86 |
| 10 | non_seasonal_expenditure | 5119.38 | 10003.86 |
| 11 | non_seasonal_frequency | 18.24 | 22.52 |
| 12 | non_seasonal_areas_dining | 0.71 | 0.46 |
| 13 | non_seasonal_areas_marina | 0.31 | 0.46 |
| 14 | non_seasonal_areas_golf | 0.31 | 0.46 |
| 15 | christmas_donation | 0.36 | 0.48 |
| 16 | scholarship_donation | 0.16 | 0.37 |
| 17 | yom | 20.11 | 11.58 |
| 18 | dist_km | 144.19 | 312.08 |

customer over the entire duration of the study. In addition, we also compute average expenditure, that is total expenditure divided by total frequency. For transaction timing, we use the average interpurchase time (*avg_ipt*) to determine how active a customer is (Reinartz & Kumar, 2003). This covariate represents the average number of days between two subsequent transactions. Specifically, we compute the difference in days between two transactions and aggregate this difference over all transactions. Finally, we divide the total difference in days by the total frequency. Note that this is not the same as the inverse of frequency, as it excludes the time between the final transaction and time at the end of study.

Furthermore, the firm's operations were subdivided into smaller business segments, such as bars or golf-shops, that represent specific areas within the firm in which a customer transactions occurred. The industry partners identified three key areas of expenditure within the firm: dining, golfing, and marina. We classify a transaction as *areas_dining* if the transaction occurs within the bar, patio, galley, main dining room, bistro, or the banquet. Similarly, we classify a transaction as *areas_golf* if the transaction occurs within the golf shop or banquet golf bar. Finally, we classify a transaction as *areas_marina* if the transaction occurs within the marina or dock shop. These indicator variables serve as proxies to measure cross-buying tendencies of a customer, which have been used in existing literature to establish relationships between the customer and the firm (Kamakura et al., 2003; Reinartz & Kumar, 2003). Note that, when a customer shows purchase activities in multiple areas of the firm, such as golfing, marina, and dining, it indicates high cross-buying tendency for the customer. An alternative to the indicator variables can be including the total expenditure in each of the business segments. An alternative model is developed using the business segment-based expenditure and the results are included in Appendix C.

We also consider donations made by a customer. We observe that some customers donate to the firm's scholarship and Christmas funds, which we denote as *scholarship_donation* and *christmas_donation* indicator variables. Similar to the cross-buying tendency variables, we include total donation amounts in an alternative model, the results of which are discussed in Appendix C.

Moreover, we include two demographic covariates: years of membership and distance of the customer from the firm in kilometres. Table 2 shows the mean and standard deviation of all covariates used in the final model computed from beginning of 2009 to end of 2016. For completeness, we also present the correlations between the covariates in Appendix A.

### 3.1. Data pre-processing

The raw data consists of transaction records for 1,368 customers from the beginning of 2009 to end of 2016. All customers in this study joined prior to 2009. Out of the 1368 customers, 1068 were active customers by the end of 2016, 177 customers resigned, and 123 customers went on a leave of absence. From the raw data, we created a database of cross-sectional data set for all customers with covariates in Table 1.

For most covariates, the standard deviations of the behavioural and demographic variables are higher than the mean. This pattern holds for active and inactive customers. The mean and standard deviation for some covariates in Table 2, such as expenditure and frequency, reveal that the covariates are highly skewed to the right. Furthermore, the correlation matrix presented in Appendix A does not show high correlation between any of the covariates.

## 4. Methods

We model customers taking a leave of absence (LOA) or resigning using a competing risk methodology. Similar to literature, here competing risk is defined as a competing event that prevents another event of interest from occurring (Gooley et al., 1999). We define the LOA event as the event of interest and the event of resigning as the competing event. In other words, a customer resigning cannot take a leave of absence. The churn probabilities of a customer facing these competing events of churn is computed by considering the absolute probability of churning from a specific event, taking into account all other competing events.

### 4.1. Competing risk survival model

Conventionally, two models have been proposed to compute risk of churn from a specific event: cumulative cause-specific hazard function (CSHF) and cumulative incidence function (CIF). Although the CSHF allows for the computation of the rate of occurrence of churn from an event of interest, it fails to account for other competing events simultaneously. For a more distinct interpretation of the risk arising from a specific event, we utilise CIF. The CIF computes the risks from the event of interest while simultaneously accounting for other competing events. However, to establish the definition of CIF, we will define the status of a customer, the CSHF estimate, and the all-cause survival.

#### 4.1.1. Customer status

The status of customer $i$ represents whether or not she faced a churn event by the end of the observation period. Following the framework of Kalbfleisch and Prentice (2011), let $\Delta_i$ represent the status of customer $i$, which can be expressed as

$$\Delta_i = \begin{cases} 0 & \text{if } C_i > T_i, \\ j & \text{if } C_i \leq T_i, \end{cases} \tag{1}$$

where $C_i$ and $T_i$ denote the censoring time and event time (i.e. time until leave of absence or resigning from the beginning of the study), respectively, $j$ denotes the specific event of churn, and 0 reflects an active customer. In other words, customer $i$ is censored if she is active at the end of the study. Alternatively, the customer is in status $j$ when the event is known by the end of the observation period (i.e. censoring time is less than event time). This customer status is employed within the CSHF.

#### 4.1.2. Competing risk hazard and survival

The CSHF function computes the instantaneous risk of churn from event $j$ for customer $i$ that is not facing any other event at time $t$. The CSHF can be expressed as

$$\lambda_j(t|\mathbf{x}_i^k) = \lim_{\Delta t \to 0} \frac{P(t \leq T_i \leq t + \Delta t, \Delta_i = j | T_i \geq t, \mathbf{x}_i^k)}{\Delta t} = \frac{f_j(t|\mathbf{x}_i^k)}{S(t|\mathbf{x}_i^k)}. \tag{2}$$

The numerator of Equation (2) is interpreted as the conditional probability of event $j$ occurring for an individual $i$ within the infinitesimal time interval $[t, t + \Delta t)$ and not before, given a vector of $k = 1, \ldots, K$ covariates $\mathbf{x}_i^k$ (Kalbfleisch & Prentice, 2011). Then, the cumulative CSHF is $H_j(t|\mathbf{x}_i^k) = \sum_t \lambda_j(t|\mathbf{x}_i^k)$. For simplicity, we exclude denoting $k$ henceforth. Finally, the sum of the independent CSHFs contributes to the cumulative all-cause hazard, which is expressed as

$$\Lambda_{.}(t|\mathbf{x}_i) = \sum_j H_j(t|\mathbf{x}_i). \tag{3}$$

The all-cause hazard function is used to compute the all-cause survival model for customer $i$ via,

$$S_{.}(t|\mathbf{x}_i) = e^{\{-\Lambda_{.}(t|\mathbf{x}_i)\}}, \tag{4}$$

which represents the probability of retention (i.e. opposite of churn) from all events at time $t$. For ease of interpretation of the individual event-specific risks, we compute the cumulative incidence estimates based on the CSHF.

#### 4.1.3. Cumulative incidence

The CIF estimates the proportion of customers at time $t$ who churned from event $j$, while accounting for customers facing other competing events. The CIF for customer $i$ facing risk $j$ can be computed via,

$$F_j(t|\mathbf{x}_i) = \sum_t \lambda_j(t|\mathbf{x}_i) S_{.}(t|\mathbf{x}_i) \qquad \forall j \in J \tag{5}$$

where $\lambda_j(t|\mathbf{x}_i)$ is the CSHF from Equation (2), and $S_{.}(t|\mathbf{x}_i)$ is the all-cause survival from Equation (4) prior to time $t$. The CIF reflects the event-specific survival, where a customer experiencing event $j$ at time $t$ must have survived all events until time $t$. In other words, the CIF demonstrates the risk of a specific event in the presence of the other events. Then, the cumulative CIF is employed in the random survival forest for the estimation of risk of churn.

### 4.2. The ensemble competing risk survival model

We employ the RSF method as it has several advantages in modelling this data, including the ability to model nonlinear effects and a lack of distributional assumptions (Ishwaran et al., 2008). RSF grows several trees in order to estimate the response. Here, the cumulative CIFs are averaged across the trees in the forest to compute the corresponding ensemble estimates. Therefore, the ensemble cumulative CIF for customer $i$ is averaged across $B$ survival trees via,

$$\bar{F}_j(t|\mathbf{x}_i) = \frac{1}{|B|} \sum_b F_{j,b}(t|\mathbf{x}_i), \tag{6}$$

where $F_{j,b}(t|\mathbf{x}_i)$ is the CIF estimate from tree $b$. This ensemble estimate is the response of the competing risk random survival forest (CR-RSF) model.

### 4.3. Prediction error

Various techniques, such as the Brier score, logarithmic scoring rule, or C-index, can be used to assess the accuracy of the risk predictions. In this work, we employ a time dependent Brier score that is evaluated using the bootstrapped cross-validation procedure (Mogensen et al., 2012). Let $D_N$ represent the entire data set with $N$ number of customers and $D_N^n$ represent a boostrapped training data set, where $N$ observations are drawn with replacement from $D_N$ with $n$ unique customers. Then, the customers not included in the training set, i.e. $D_N/D_n$, serves as the test set. For example, if the entire dataset consists of $D_N = 50$ customers, then the training set can be also comprised of $D_N^n = 50$ with $n = 20$ unique customers. Then, the remaining $D_N/D_n = 30$ customers make up the testing set. Finally, the bootstrapping is repeated $B$ times. The bootstrap cross validation error is computed via,

$$\text{BootCvErr}(t) = \frac{1}{B}\sum_{b=1}^{B}\sum_{i\in D_N/D_n}\omega_i(t)\{X_{i,j}(t)-\bar{F}_{i,j,s}(t|\mathbf{x}_i)\}^2. \tag{7}$$

where,

$$X_{i,j}(t) = \begin{cases} 1 & \text{if customer } i \text{ experiences event } j \text{ before time } t \\ 0 & \text{if customer } i \text{ is event-free or faces a competing event before time } t \end{cases}$$

and $\bar{F}_{i,j,s}(t|\mathbf{x}_i)$ is the CIF of customer $i$ who experienced event $j$, and $\omega_i(t)$ are the inverse probability censoring weights. Since survival status at time $t$ could also be right censored, the inverse probability censoring weights help avoid bias in the population average (Gerds & Schumacher, 2006; Graf et al., 1999). We can summarise the prediction error computed from Equation (6) using the integrated Brier score (IBS), which is a cumulative Brier score, measured from a minimum (0 days) to a maximum (7 years in this data) time period.

### 4.4. Model interpretation

Although machine learning is better suited for prediction rather than inference, results from random forests do provide some level of interpretation. Tools such as variable importance (VI) and partial dependence plots (PDP) provide insights that help understand customer behaviour in relation to outcomes of churn.

**Variable importance** is used to identify the more salient covariates by assessing the degree to which the covariate is accurate in predicting the outcome. Let $\mathbf{x}_i$ denote the vector of $k$ covariates for customer $i$. Then the VI computation comprises of four steps.

Step 1: Randomly permute the values of a covariate $k$ in vector $\mathbf{x}_i^k$. The random permutations help capture the level of association between the covariate and the outcome. Denote the permuted vector $\mathbf{x}_i^{k'}$ where covariate $k$'s value is replaced by the permuted value.

Step 2: Use the vector of covariates $\mathbf{x}_i^{k'}$ and the vector of covariates $\mathbf{x}_i^k$ to predict the outcomes $\hat{y}'_{OOB,i}$ and $\hat{y}_{OOB,i}$, respectively, from the out-of-bag (OOB) samples in the forest. Here, OOB is the testing sample for computing errors.

Step 3: Compute the OOB-error rates by taking the difference in the prediction and actual values, $\hat{y}'_{OOB,i} - y_{OOB,i}$ for the permuted covariate and $\hat{y}_{OOB,i} - y_{OOB,i}$ for the non-permuted covariate.

Step 4: Average the difference between the error rates over the forest, which is referred to as the permutation variable importance (Breiman, 2001).

If the random permutation causes a large difference in average error rates in step 4, it implies that covariate $k$ is significantly associated with the response. Consequently, higher error indicates higher importance of the covariate while a low or negative VI indicates low importance or noise. An added advantage of VI is that it can accommodate the different outcomes of churn, as in this case, by computing event-specific variable importances (Ishwaran et al., 2010).

**Partial dependence** is used to understand the relationships between the predicted outcome and a covariate of interest for tree-based models. To compute the partial dependence for covariate $k$, let $\mathbf{x}'_i$ be the vector of covariates that excludes covariate $k$ for customer $i$. Also, let $\hat{f}(\mathbf{x})$ be the unknown functional form estimated by the forest. Then, the partial dependence is given by

$$f(x_i^k) = E_{\mathbf{x}'_i}\{\hat{f}(x_i^k, \mathbf{x}'_i)\} = \int \hat{f}(x_i^k, \mathbf{x}'_i)p(\mathbf{x}'_i)d\mathbf{x}'_i \tag{8}$$

where $p(\mathbf{x}'_i)$ is the marginal probability density of $\mathbf{x}'_i$ (Greenwell, 2017).

## 5. Results

In this section, we first establish a decision-criteria to classify the status of a customer using the predicted risk of churn. Second, based on the classifications, we (i) identify important covariates that contribute to a specific churn outcome in subsection 5.2.1 and (ii) explore covariate-churn relationships for the important covariates in subsection 5.2.2. Furthermore, we compare the direction of impact of important covariates and the prediction accuracy to

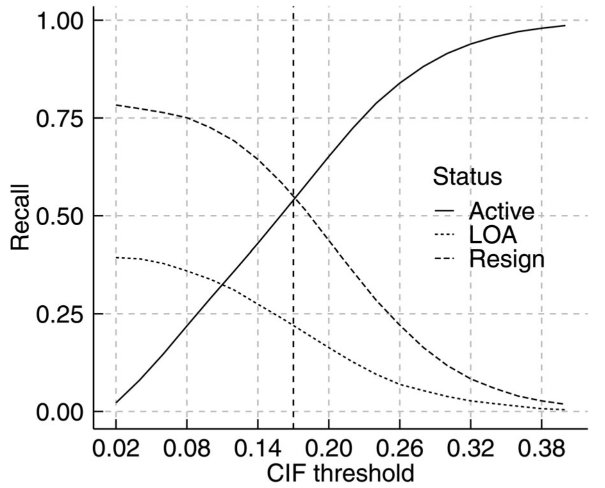



Figure 1. Recall for different threshold CIFs for the event of active, LOA or resign.

existing parametric methods in churn management in subsection 5.3. Moreover, we also compare the accuracy of the proposed method with a traditional multinomial logistic regression model, the results of which are reported in Appendix D.

### 5.1. Customer status

We use CIF computed by CR-RSF to predict the status of a customer (i.e. active, LOA or resign). To define a status, we first establish a threshold that produces the highest accuracy across the three statuses. Then, if a customer's CIF for LOA and resign is lower than the CIF threshold, then the customer's predicted status is considered to be active. If a customer's estimated CIF for either LOA or resign is higher than the CIF threshold, then the customer's predicted status is considered to be either LOA or resigned, depending on whichever CIF is higher. Then, we construct a $3 \times 3$ confusion matrix of the predicted versus actual statuses using bootstrap cross-validation as outlined in Subsection 4.3. These confusion matrices are used to compute recall. Additionally, precision and F1 score are computed and shown in Appendix B (Hastie et al., 2001).

Based on Figure 1, we employ a threshold CIF of 0.17 as it places equal weight to the recall scores for active, LOA and resign. However, a manager may chose a threshold CIF depending on the outcome that needs to be prioritised. For instance, if accuracy of churn events are of higher importance, the manager may choose a threshold of 0.14. On the other hand, if accuracy of predicting active is preferred, then the manager may choose a threshold of 0.26.

### 5.2. Customer behaviour

After classifying customers, we explore the relationships between their behaviour and risks of churn. Specifically, we first identify important behaviours that contribute to churn using variable importance. Then, we explore the relationships between important covariates and the risk from a competing event using the PDP.

#### 5.2.1. Important behaviours

Higher values of VI indicate more important covariates. In practice, VI $\approx 0$ or VI $< 0$ are considered to be unimportant for prediction and often excluded from analysis. Here, we observe both donation covariates along with *seasonal_areas_marina*, *non_seasonal_areas_marina*, and *non_seasonal_areas_golf* have negligible importance (not shown in Figure 2). Therefore, we create a reduced forest by omitting these less important covariates and report the results of the reduced CR-RSF, henceforth.

Figure 2 shows *seasonal_areas_dining* is the most important covariate for both events, which means a customer's activity in dining areas of the firm during peak-season highly contributes to the prediction accuracy of overall churn. Note that, for both events of churn, purchase activity in the other areas of the firm, golfing and marina, is not important. This demonstrates that an average customer's cross-buying tendency during peak-season is not important for churn prediction. Rather covariates related to transaction intensity and transaction timing is more important for predicting churn, as observed in existing studies (Reinartz et al., 2005). Additionally, covariates related to time between purchases and number of purchases are more important for both events of churn, relative to covariates based on the dollar amount of purchases, such as expenditure. Years of membership is more important for leave of absence than for resign. In other words, a customer's tenure in the firm plays an important role in determining if they will go on a leave of absence.

The VI plot allows marketers to (i) determine the relative importance of a covariate within a specific event of churn and design event-specific incentives to influence a particular covariate and (ii) determine the relative impact of an event on a covariate of interest. For the event of LOA, *seasonal_areas_dining* is the most important variable followed by *yom*, *seasonal_frequency*, and *seasonal_avg_ipt*. The presence of peak-season covariates in the top four indicates that a decline in a customer's purchase activity during peak-seasons is an important determinant of a customer's propensity to temporarily disengage her relationship with the firm. For resign, there are seasonal and off-season transaction timing-based covariates among the top four. The presence of off-season and peak-season covariates indicate that monitoring a customer's activity during the entire year is important for determining a customer's

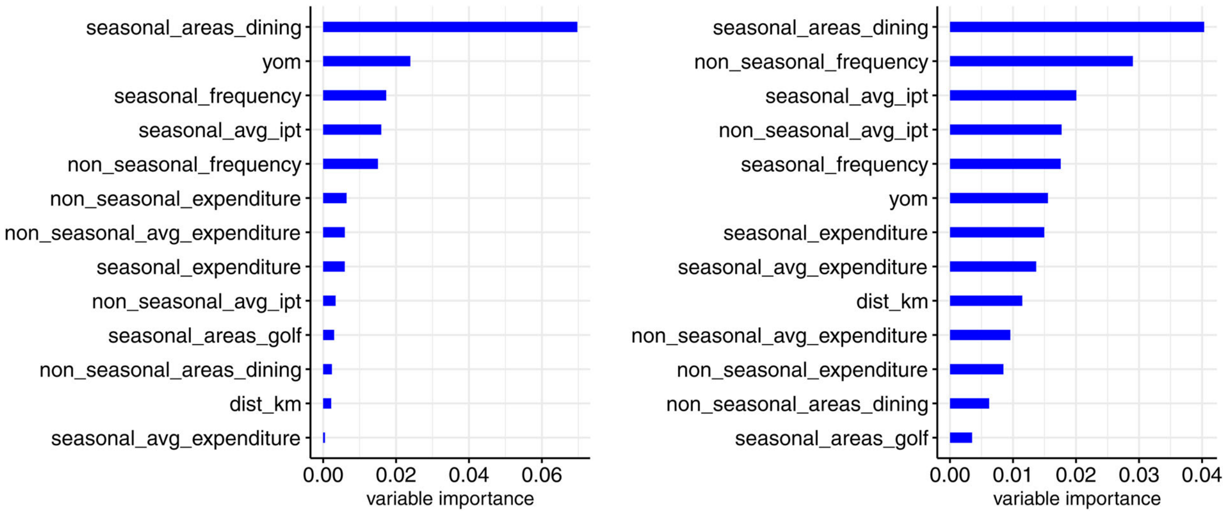



**Figure 2.** Variable importance of covariates for the event of leave of absence (left) and resign (right).

propensity to resign. Consequently, in order to make decisions on how to provide incentives for customers predicted to resign, a marketer must consider a larger period for customer's purchase history. However, a shorter time-frame can be used to model for leave of absence. If such differences between the two churn events are ignored, it could result in sub-optimal churn management. Given limited resources, VI can allow decision-makers to prioritise in monitoring important covariates compared to less important ones. Furthermore, marketers can combine the importance of both events to compute an overall importance of variables to affecting risk of churn. This can help marketers monitor possibly a lesser set of covariates that simply affect overall churn.

#### *5.2.2. Behaviour-churn relationships*

We explore the relationships between churn and the important behaviours identified by VI using partial dependence plots. Specifically, we explore the overall functional relationship between a covariate of interest and a churn event. Figure 3 and Figure 4 show the partial dependence plots for the important variables for the event of LOA and resign, respectively. For indicator variables, partial dependence boxplots are employed.

In the context of leave of absence, the boxplot PDP for *seasonal_areas_dining* shown in Figure 4 suggest that risk of LOA is higher for customers that spend in dining areas within the firm during peak-season. This suggests that there are customers who spend in dining areas during peak-season with the intention of using the firm's services only during that time-period, and going on a leave of absence during off-season. The PDP for years of membership suggests that risk of LOA reduces as the tenure of the customer increases. This suggests that, in general, that customers with longer tenure are loyal customers for this firm, which is also confirmed by the PDP for resign. The PDP for peak-season frequency show an increasing and then plateauing relationship with the risk of LOA. Although this effect may seem inverse at first, the increase can be attributed to customers who make frequent smaller transactions during the peak-season (such as the dining services) with the intention of only using it temporarily. For off-season frequency, there exists a desirable frequency that minimises LOA risk. Customers who transact more or less are at the risk of LOA. For seasonal interpurchase times, the risk from LOA increases and then plateaus, similar to peak-season frequency. This effect shows that there are some customers who have much lower IPT and transact quite frequently and remain loyal. However, after a certain threshold such as 20, the risk remains unchanged. In summary, PDP reveals a negative linear relationship for years of membership and non-linear relationships for peak-season frequency, peak-season average IPT and off-season frequency.

In the context of resign, the boxplot PDP in Figure 4 also suggests that risk of resign is higher for customers who do spend in dining areas during peak-season. However, this increase in risk is much smaller compared to LOA. This shows the dining services attract customers who have the intention of using the membership temporarily. PDP for off-season frequency indicates that risk increases and then decreases as frequency increases. In other words, customers who transact infrequently or very frequently during off-season will not resign. However, some customers who transact around 20–25 times are at risk of resigning, opposite of LOA. For peak-season interpurchase time, the risk of resign increases as as time increases, similar to LOA. Similar increasing trends are observed between average off-season inter-purchase time and risk. The risk of resign has a desirable value for peak-season

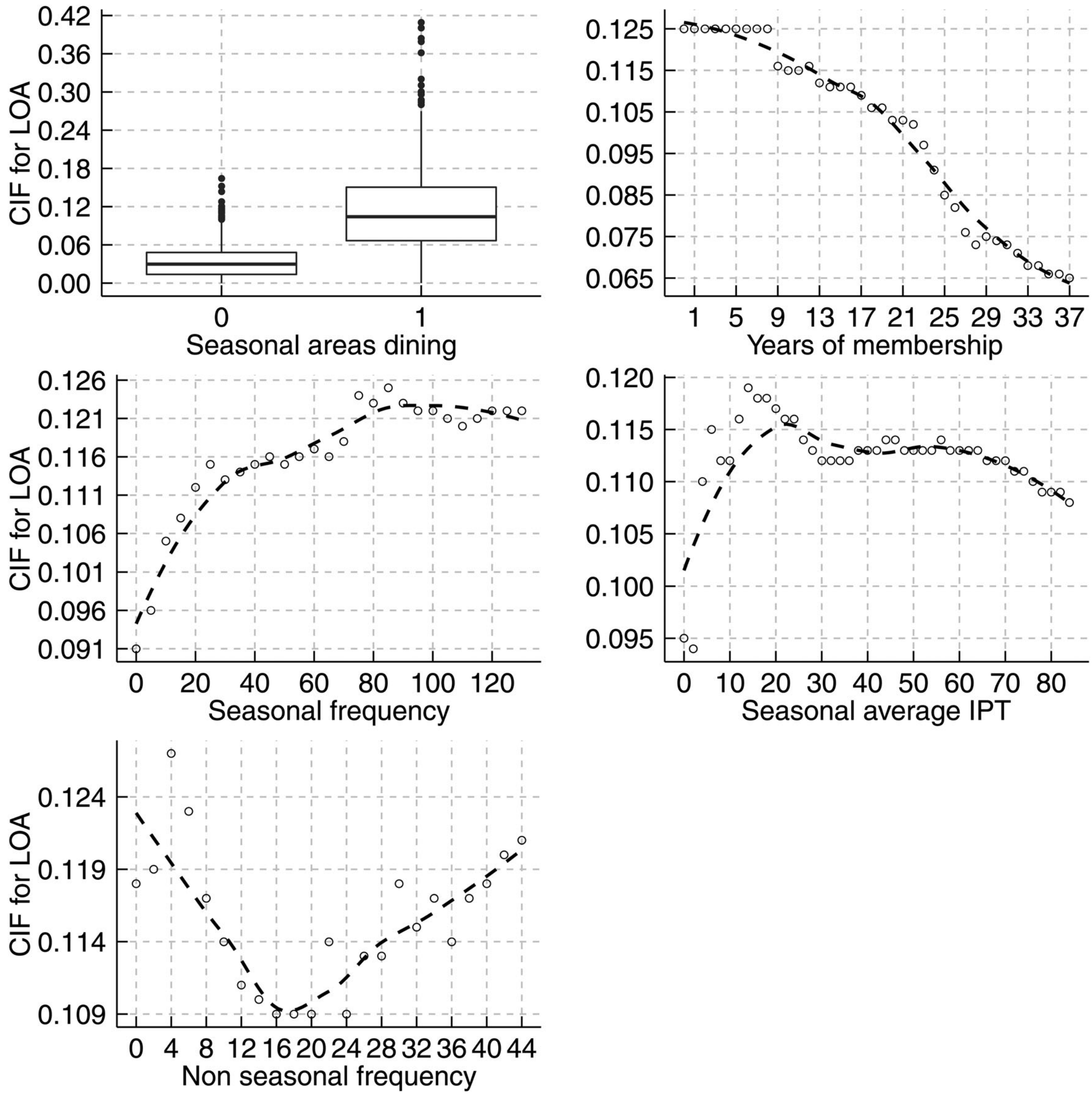



**Figure 3.** PDP for important covariates for leave of absence. Note that for continuous covariates, we compute the CIF values for covariate values corresponding to the top 90th percentile.

frequency, similar to off-season frequency for LOA. This effect may be an artefact of the increased loyalty from customers with longer tenure. The trends for peak-season total and average expenditures are similar with a negative effect on resigning. Years of membership shows a steady decrease in risk of resigning. However, for years of membership above 30, there is an increase, which could suggests the end of tenure for older customers who either are unable to use the services due to health issues or have passed away. In summary, here, the PDP reveals linear relationships for off-season average IPT and peak-season total and average expenditures, and nonlinear effects for peak- and off-season frequency, peak-season average IPT and years of membership.

While most of the relationships could be explained by behavioural phenomenons, the relationship between peak-season frequency and the estimated risk of LOA is rather unusual. Ideally, LOA risk should decrease for customers with higher frequency of purchase (Reinartz & Kumar, 2003). Here, we suspect this result to be an effect of an interaction between the frequency of purchase and years of membership. Therefore, we employ a conditional dependence plot (CDP) to interpret the interaction between these two variables. For tree based models such as CR-RSF, a CDP is a visualising tool for plotting the response across levels of two (or more) covariates (Cleveland, 1993). Here, we construct a CDP by grouping the partial dependencies of years of memberships for five different ranges of frequencies as shown in Figure 5.

The CDP plot reveals that risk of LOA increases for customers with transaction frequency ranges between (0,41] to (41,82]. However, following this unexpected increase, the risk of leaving declines as peak-season frequency of purchases rise, which is aligned with existing literature (Reinartz & Kumar, 2003). The increase in risk can be attributed to the

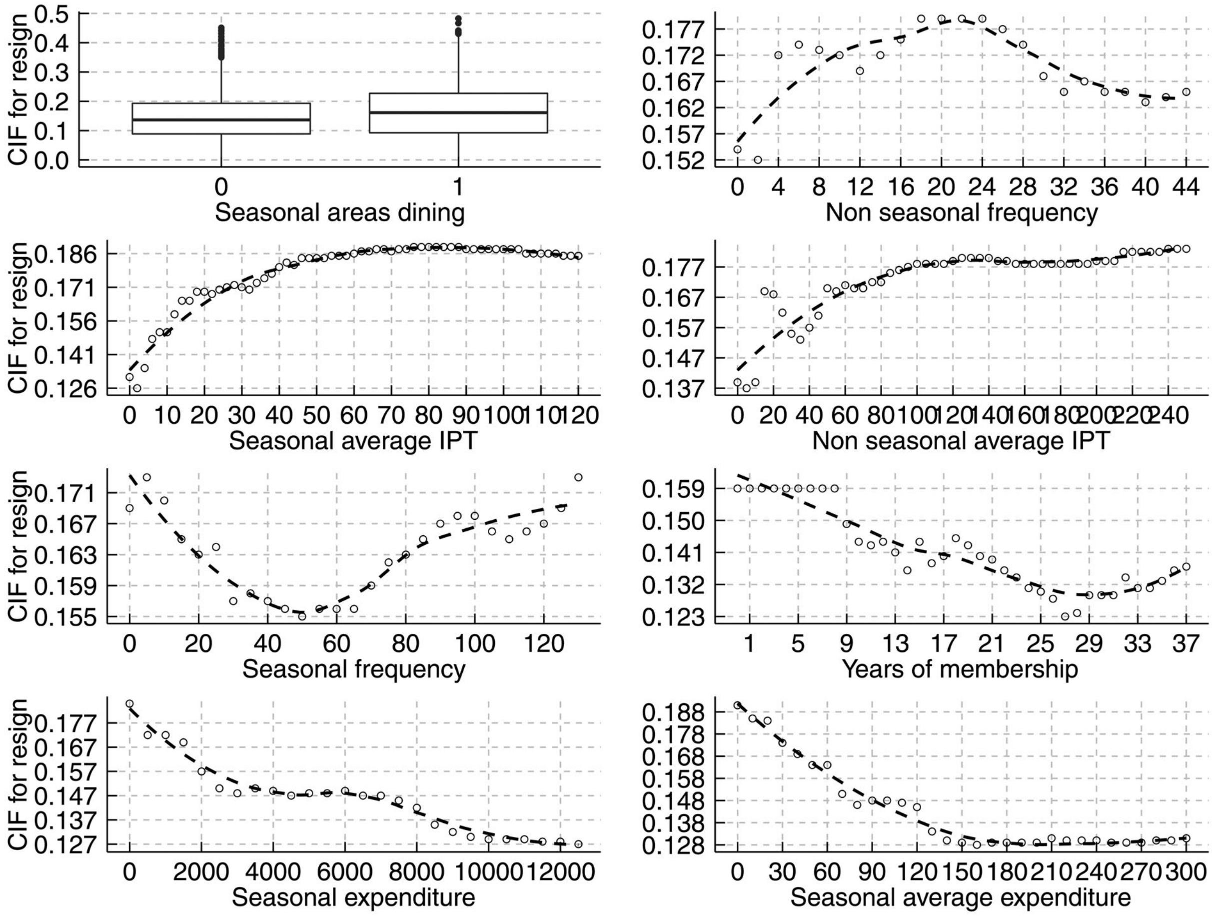



**Figure 4.** PDP for important covariates for resign. Note that for continuous covariates, we compute the CIF values for covariate values corresponding to the top 90th percentile.

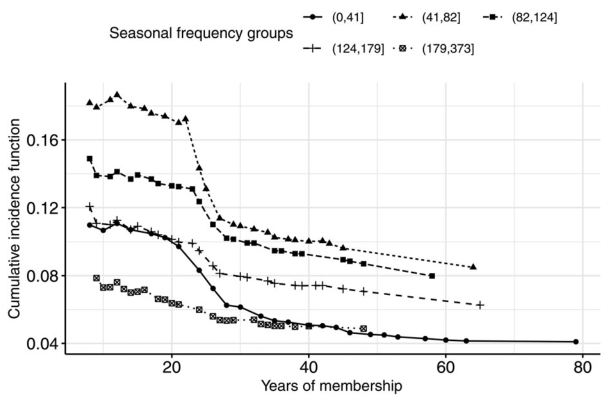



**Figure 5.** Conditional dependence plot of frequency across years of membership.

customer's years of membership within each group. Figure 5 shows that more of the senior customers (who have been with the firm longer than 40 years) had less frequent purchases, within the (0,41] range. The effect of years of membership, therefore, plays a role in decreasing the values for CIF within this *seasonal_frequency* group.

### 5.3. Model performance

To validate our findings and measure performance, we compare (i) the direction of impact from the functional relationships observed in the PDPs, and (ii) the predictive accuracy of our CR-RSF against traditional econometric models. In this data, we observe that 100% of customers who opt for LOA and 92% of customers who opt for resign have transactions in *areas_dining* during peak-seasons. This leads to complete separation for the event of LOA for people who have transactions in dining areas during peak-season. Complete separation is a limitation for econometric models. While the CR-RSF model overcomes this limitation, we remove *seasonal_areas_dining* from all models (including CR-RSF) for a consistent comparison.

#### 5.3.1. Direction of impact

The PDPs in Figures 3 and 4 demonstrate the direction of impact for the important covariates. Further, the presence of nonlinear relationships are observed between each important covariate (excluding seasonal areas dining) and each churn event for the CR-RSF. For the econometric methods, three parametric and semi parametric competing risk models are employed using the same set of important covariate for each event. The semi-parametric competing risk Cox model (CR-Cox) measures the impact of covariates on the event specific hazard rates (Cox, 1992). The Fine and Gray (FG) model measures the

**Table 3.** Model estimates and direction of impact for parametric models. **f(x)** is the transformation function applied to the parametric model and # denotes opposite direction of impact in CR-RSF PDP.

| Event | Covariate | f(x) | Model estimates | | | Direction of impact |
|---|---|---|---|---|---|---|
| | | | CR-Cox[a] | Fine & Gray[b] | CR-AFT (Weibull)[c] | |
| LOA | seasonal_avg_ipt | Log(*f*(*x*)) | 2.6884*** | 2.5822*** | −2.6528*** | Inverted U |
| | | Log(*f*(*x*)$^2$) | −0.4287*** | −0.4161*** | 0.4137*** | |
| | seasonal_frequency | Log(*f*(*x*)) | 0.8113** | 0.7916*** | −0.8155** | Positive |
| | yom | *none* | −0.0477*** | −0.0461*** | 0.0441*** | Negative |
| | non_seasonal_frequency | Log(*f*(*x*)) | 0.5090 | 0.4422 | −0.4776 | Inverted U# |
| | | Log(*f*(*x*)$^2$) | −0.1631* | −0.1485* | 0.1494* | |
| Resign | non_seasonal_frequency | Log(*f*(*x*)) | 0.9474** | 0.9328* | −0.9625** | Inverted U |
| | | Log(*f*(*x*)$^2$) | −0.2031** | −0.1933** | 0.2007** | |
| | seasonal_avg_ipt | $log(f(x))$ | 1.5181*** | 1.5170* | −1.5797* | Inverted U |
| | | Log(*f*(*x*)$^2$) | −0.1644*** | −0.1643 | 0.1727 | |
| | non_seasonal_avg_ipt | Log(*f*(*x*)) | 0.3189 | 0.3207* | −0.3315* | Positive |
| | seasonal_frequency | Log(*f*(*x*)) | 2.5458* | 2.4586** | −2.4255* | Inverted U# |
| | | Log(*f*(*x*)$^2$) | −0.0793 | −0.0719 | 0.0753 | |
| | yom | *none* | −0.0374* | −0.0322 | 0.0342 | U-shaped |
| | | *f*(*x*)$^2$ | 0.0007 | 0.0007* | −0.0007* | |
| | seasonal_expenditure | Log(*f*(*x*)) | −1.4167 | −1.4044* | 1.3227 | Negative |
| | seasonal_avg_expenditure | Log(*f*(*x*)) | 1.1119 | 1.1143 | −1.0303 | Positive# |

Significance codes: '***' 0.001 '**' 0.01 '*' 0.05.
[a]The CR-Cox and FG frameworks models the probability of risk – using cause specific cumulative hazard and cumulative incidence functions respectively – based on covariates.
[b]The Weibull AFT models event specific survival and not risk. The directions are therefore, reversed.
[c]The intercepts were significant. The value of log scale was −0.07 for LOA and -0.02 for resign, which were estimated using curve-fitting.

impact of the covariates on the CIF directly (Fine & Gray, 1999). The parametric competing risk Accelerated Failure Time (AFT) model (with an underlying Weibull distribution) measures the impact of the covariates on event specific survival (Kumar et al., 2018). Note that since CR-AFT measures survival, the direction of impact are expected to be complementary to the other models measuring risk. Additionally, we consider log transformations and squared terms for these econometric models to normalise for skewed distributions and nonlinear relationships. The functional relationship for each covariate for each econometric model and for each event is summarised in Table 3. We compare the overall direction of impact for the covariates from the PDP for the CR-RSF with the direction of impact from the coefficients for the three econometric models.

Table 3 shows that results from the PDP of CR-RSF model are largely in agreement with the directions obtained from the econometric models. With the exception of seasonal frequency and seasonal average expenditure, the results are identical for all other covariates across all models for both events. For instance, the CR-RSF PDP for interpurchase time shows an inverted U shaped effect both on the risk of resign and LOA, which is consistent with literature (Reinartz & Kumar, 2003). All econometric models show the same directional impact. The U shaped effect uncovered by the PDP for years of membership can also be validated by the econometric models. However, there are some differences. Peak-season average expenditure shows an overall negative relationship with risk of churn for the CR-RSF and a positive relationship for the parametric models. A negative relationship is expected, since higher average spending per transaction should reduce the risk of churn. In this case, the functional form of the parametric models could be mis-specified leading to an incorrect relationship. However, the CR-RSF remains to robust to such limitation and produces directions of impact that is consistent with existing literature (Reinartz et al., 2005). Overall, most covariates found to be important by CR-RSF are also significant across the parametric models, further validating the our proposed CR-RSF model.

### 5.3.2. Model performance

We employ the FG model as the benchmark for the proposed CR-RSF, as FG models use cumulative incidence as a response, similar to our CR-RSF. CR-Cox or CR-AFT models use hazard rather than CIF. Therefore, we exclude these two methods for benchmarking purposes. We compare the prediction error for different quantiles of customer segments based on their *seasonal_expenditure* values from lowest to highest. Specifically, for prediction error, IBS scores for the CR-RSF and FG are computed over 100 bootstraps according to Equation (7). The results are summarised in Figure 6.

For both events, CR-RSF outperforms the parametric FG model for the top 25%, 50%, and 75%, quantiles by reducing prediction error. The improvement is larger for more valuable customer segments, such as the top 25% of customer. This finding would also improve prediction and targeting of high-risk high-value customers. The improved accuracy demonstrates the robustness of the CR-RSF model to account for unspecified functional

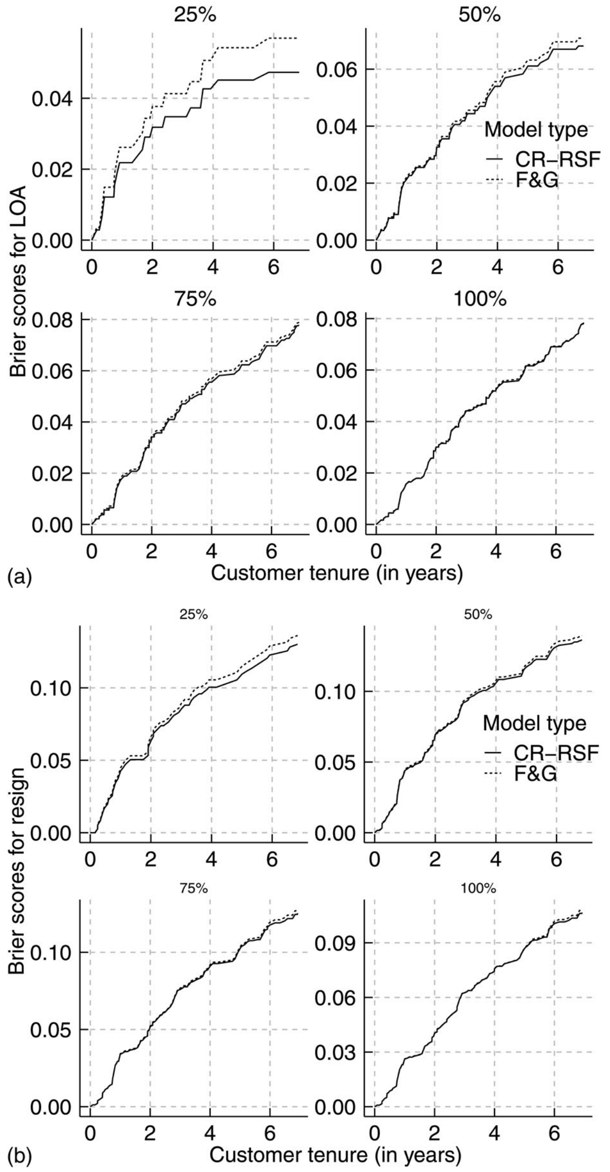



Figure 6. Left) LOA and Right) Resign: Bootstrapped cross-validated integrated Brier score (IBS) computed over 7 years for a Fine & Gray model and the proposed CR-RSF model. The IBS scores are measured for different quantiles of customer base.

forms, unlike the FG model. In addition, we observe that the prediction error increases, for both models, as we include larger ordered samples. In other words, as we consider more and more customers who have longer and longer interpurchase times, the error for all models increase.

## 6. Discussion

Churn management in the presence of uncertainties is a challenge for marketers. Here, we take a first step in addressing this challenge by identifying the churners with accuracy, and present strategies on how to monitor behaviour that can be effective in decreasing churn. Furthermore, we account for competing events of churn that exist in various membership-based industries. For instance, satellite television providers and other telecommunication firms allow customers to go on a leave of absence and pay a substantially reduced fee for the leave duration. Under these circumstances, when the firm wishes to provide incentives to curb customers from leaving or resigning, the proposed method can aid in decision-making.

In the context of research implications, this study uses a competing risk random survival forest method to estimate the churn probabilities of customers. This methodology can be readily extended to any number of competing events under a contractual business setting. Furthermore, the proposed method provides reliable results with higher accuracy without prior knowledge on the exact functional form that best describes the relationship between the dependent and independent variables. However, RSF does rely on a large quantity of available data to train the model. As a result, even greater accuracy can be achieved with larger data sets, which we have planned for in future extensions of this study.

The RSF algorithm implemented here adheres to the framework described by Breiman (2001) when incorporating the competing risks. Other extensions of RSF, such as conditional inference forests, are yet to be developed for the competing risk paradigm. Such methods could potentially improve the computation speed and accuracy of predictions. Additionally, the proposed model can be extended to incorporate the changing values of covariates over time. This modelling effort serves as an extension of this study.

Although, we include customer behavioral measures that are commonly used in literature, our list of antecedents of churn is not meant to be exhaustive. We exclude customer behavioral metrics such as recency or firm level variables such as advertising expense, number of employees, and promotional expense from our analysis. We choose to exclude recency primarily for two reasons, (i) unlike other customer behaviour metrics, recency is not aggregated over time and only captures the time since the last purchase, and (ii) to avoid the recency trap whereby customers who haven't spent for a long time may still be active (Neslin et al., 2013). Furthermore, firm level variables such as advertising expense, number of employees, and promotional expense were not readily available for this study. However, we plan to include these covariates in future work.

In the context of managerial implications, this work shows an accurate way to predict churners, in particular when facing competing risks. Further, we demonstrate how variable importance allows marketers to interpret the relative impact of covariates on the competing events of churn. Moreover, PDP and CDP can help a marketer understand the relationships between covariates and the risk of churning. Collectively, these results would guide marketers to prescribe better actions in terms of designing and implementing an effective churn management program.

Accurate churn predictions can also improve predictions of customer lifetime value (CLV), a widely used measure of customer worth in industry. Ultimately, managers can employ the CR-RSF framework to improve CLV predictions and better manage customer relationships (Routh, 2018). However, there are some limitations to the managerial implications from this study; the data used in this work is restricted to a membership-based firm in the hospitality industry. To investigate applicability of our findings, additional data from other membership-based industries can be analysed using the CR-RSF framework. Furthermore, other customer demographic covariates, such as race or gender, or behavioural covariates, such as branch switching tendencies, can be considered for future studies. The effects of external factors, such as market size and the extent of competition, can also be considered for future work. Moreover, the proposed method is suited for a contractual business setting where the status of the customer is known. Further research can be done to investigate similar goals under non-contractual business settings.

## 7. Conclusion

Predicting customer churn with accuracy is essential to churn management. Additionally, understanding how customer behaviour contributes to customer churn is important in designing the churn management incentives. Uncertainties, such as competing events of churn, and mis-specifications in prediction models often limit the success of churn management programs.

We demonstrate a competing risk random survival forest model to estimate the churn probabilities of a customer within a contractual business setting in the presence of competing events of churn. The method outperforms conventional methods by providing higher accuracy in identifying potential churners. Furthermore, the model does not rely on a specific functional form to model risk of churn based on customer behaviour. We validated the findings from this model by comparing the results to existing econometric models. This framework for evaluating customer churn under competing risk can be readily incorporated in the real world due to predictive power and lack of assumptions for improved decision-making.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

Pallav Routh http://orcid.org/0000-0002-6920-527X
Arkajyoti Roy http://orcid.org/0000-0003-0348-9297
Jeff Meyer http://orcid.org/0000-0001-7827-8326

## References

Ascarza, E., Iyengar, R., & Schleicher, M. (2016). The perils of proactive churn prevention using plan recommendations: Evidence from a field experiment. *Journal of Marketing Research*, *53*(1), 46–60. https://doi.org/10.1509/jmr.13.0483

Ascarza, E., Neslin, S. A., Netzer, O., Anderson, Z., Fader, P. S., Gupta, S., Hardie, B. G. S., Lemmens, A., Libai, B., Neal, D., Provost, F., & Schrift, R. (2018). In pursuit of enhanced customer retention management: Review, key issues, and future directions. *Customer Needs and Solutions*, *5*(1-2), 65–81. https://doi.org/10.1007/s40547-017-0080-0

Ascarza, E., Netzer, O., & Hardie, B. G. (2018). Some customers would rather leave without saying goodbye. *Marketing Science*, *37*(1), 54–77. https://doi.org/10.1287/mksc.2017.1057

Bodapati, A., & Gupta, S. (2004). A direct approach to predicting discretized response in target marketing. *Journal of Marketing Research*, *41*(1), 73–85. https://doi.org/10.1509/jmkr.41.1.73.25081

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. https://doi.org/10.1287/mksc.17.1.45

Braun, M., & Schweidel, D. A. (2011). Modeling customer lifetimes with multiple causes of churn. *Marketing Science*, *30*(5), 881–902. https://doi.org/10.1287/mksc.1110.0665

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *24*(2), 123–140. https://doi.org/10.1007/BF00058655

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Cleveland, W. S. (1993). *Visualizing data. visualizing data. visualizing data*. Hobart Press.

Cox, D. R. (1992). Regression models and life-tables. In *Breakthroughs in statistics* (pp. 527–541). Springer.

Dudoit, S., Fridlyand, J., & Speed, T. P. (2002). Comparison of discrimination methods for the classification of tumors using gene expression data. *Journal of the American Statistical Association*, *97*(457), 77–87. https://doi.org/10.1198/016214502753479248

Fine, J. P., & Gray, R. J. (1999). A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association*, *94*(446), 496–509. https://doi.org/10.1080/01621459.1999.10474144

Gerds, T. A., & Schumacher, M. (2006). Consistent estimation of the expected brier score in general survival models with right-censored event times. *Biometrical Journal. Biometrische Zeitschrift*, *48*(6), 1029–1040. https://doi.org/10.1002/bimj.200610301

Gönül, F. F., Kim, B.-D., & Shi, M. (2000). Mailing smarter to catalog customers. *Journal of Interactive Marketing*, *14*(2), 2–16. https://doi.org/10.1002/(SICI)1520-6653(200021)14:2<2::AID-DIR1>3.0.CO;2-N

Gooley, T. A., Leisenring, W., Crowley, J., & Storer, B. E. (1999). Estimation of failure probabilities in the presence of competing risks: New representations of old estimators. *Statistics in Medicine*, *18*(6), 695–706. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O

Graf, E., Schmoor, C., Sauerbrei, W., & Schumacher, M. (1999). Assessment and comparison of prognostic classification schemes for survival data. *Statistics in Medicine*, *18*(17-18), 2529–2545. https://doi.org/10.1002/(SICI)1097-0258(19990915/30)18:17/18<2529::AID-SIM274>3.0.CO;2-5

Greenwell, B. M. (2017). pdp: An r package for constructing partial dependence plots. *The R Journal*, *9*(1), 421–436. https://doi.org/10.32614/RJ-2017-016

Gupta, S., Lehmann, D. R., & Stuart, J. A. (2004). Valuing customers. *Journal of Marketing Research*, *41*(1), 7–18. https://doi.org/10.1509/jmkr.41.1.7.25084

Gustafsson, A., Johnson, M. D., & Roos, I. (2005). The effects of customer satisfaction, relationship commitment dimensions, and triggers on customer retention. *Journal of Marketing*, *69*(4), 210–218. https://doi.org/10.1509/jmkg.2005.69.4.210

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.

Hoel, D. G. (1972). A representation of mortality data by competing risks. *Biometrics*, *28*(2), 475–488. https://doi.org/10.2307/2556161

Ishwaran, H., Gerds, T. A., Kogalur, U. B., Moore, R. D., Gange, S. J., & Lau, B. M. (2014). Random survival forests for competing risks. *Biostatistics (Oxford, England)*, *15*(4), 757–773. https://doi.org/10.1093/biostatistics/kxu010

Ishwaran, H., Kogalur, U. B., Blackstone, E. H., & Lauer, M. S. (2008). Random survival forests. *The Annals of Applied Statistics*, *2*(3), 841–860. https://doi.org/10.1214/08-AOAS169

Ishwaran, H., Kogalur, U. B., Gorodeski, E. Z., Minn, A. J., & Lauer, M. S. (2010). High-dimensional variable selection for survival data. *Journal of the American Statistical Association*, *105*(489), 205–217. https://doi.org/10.1198/jasa.2009.tm08622

Kalbfleisch, J. D., & Prentice, R. L. (2011). *The statistical analysis of failure time data* (Vol. 360). John Wiley & Sons.

Kamakura, W. A., Wedel, M., De Rosa, F., & Mazzon, J. A. (2003). Cross-selling through database marketing: A mixed data factor analyzer for data augmentation and prediction. *International Journal of Research in Marketing*, *20*(1), 45–65. https://doi.org/10.1016/S0167-8116(02)00121-0

Knott, A., Hayes, A., & Neslin, S. A. (2002). Next-product-to-buy models for cross-selling applications. *Journal of Interactive Marketing*, *16*(3), 59–75. https://doi.org/10.1002/dir.10038

Kumar, V., Leszkiewicz, A., & Herbst, A. (2018). Are you back for good or still shopping around? investigating customers' repeat churn behavior. *Journal of Marketing Research*, *55*(2), 208–225. https://doi.org/10.1509/jmr.16.0623

LeBlanc, M., & Crowley, J. (1995). A review of tree-based prognostic models. In *Recent advances in clinical trial design and analysis* (pp. 113–124). Springer.

Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, *43*(2), 276–286. https://doi.org/10.1509/jmkr.43.2.276

Levinthal, D. A., & Fichman, M. (1988). Dynamics of interorganizational attachments: Auditor-client relationships. *Administrative Science Quarterly*, *33*(3), 345–369. https://doi.org/10.2307/2392713

Lewis, M. (2004). The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research*, *41*(3), 281–292. https://doi.org/10.1509/jmkr.41.3.281.35986

Menard, S. (2002). *Applied logistic regression analysis* (Vol. 106). Sage.

Mogensen, U. B., Ishwaran, H., & Gerds, T. A. (2012). Evaluating random forests for survival analysis using prediction error curves. *Journal of Statistical Software*, *50*(11), 1–23. https://doi.org/10.18637/jss.v050.i11

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. *Journal of Marketing Research*, *43*(2), 204–211. https://doi.org/10.1509/jmkr.43.2.204

Neslin, S. A., Taylor, G. A., Grantham, K. D., & McNeil, K. R. (2013). Overcoming the "recency trap" in customer relationship management. *Journal of the Academy of Marketing Science*, *41*(3), 320–337. https://doi.org/10.1007/s11747-012-0312-7

Prentice, R. L., & Gloeckler, L. A. (1978). Regression analysis of grouped survival data with application to breast cancer data. *Biometrics*, *34*(1), 57–67. https://doi.org/10.2307/2529588

Reinartz, W., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. *Journal of Marketing*, *67*(1), 77–99. https://doi.org/10.1509/jmkg.67.1.77.18589

Reinartz, W., Thomas, J. S., & Kumar, V. (2005). Balancing acquisition and retention resources to maximize customer profitability. *Journal of Marketing*, *69*(1), 63–79. https://doi.org/10.1509/jmkg.69.1.63.55511

Routh, P. (2018). *A framework for estimating customer worth under competing risks*. (Unpublished doctoral dissertation), Bowling Green State University.

Venkatesan, R., & Kumar, V. (2004). A customer lifetime value framework for customer selection and resource allocation strategy. *Journal of Marketing*, *68*(4), 106–125. https://doi.org/10.1509/jmkg.68.4.106.42728

Wedel, M., & Kannan, P. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(6), 97–121. https://doi.org/10.1509/jm.15.0413

## Appendix A covariate correlations

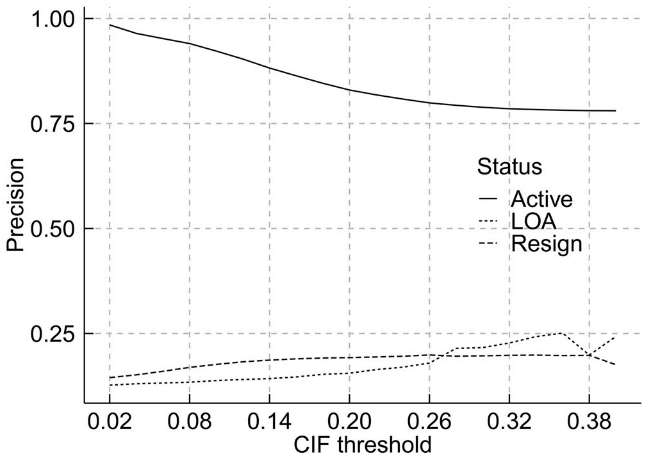
See Figure 1A.

## Appendix B precision and F1 score for CR-RSF

The measures for precision and F1 score is shown in Figure B1. The area under the curve for precision is not as high as recall for the two competing events across all possible thresholds. However, the precision plots were better for predicting the accuracy of active customers. For a threshold of 0.17, for instance, the precision scores are lower in comparison to recall for the competing events of churn, but they are higher for the active customers. For F1 score, a threshold of 0.14 focuses on higher accuracy of churners, or a threshold value of 0.26 on higher

| | non_seasonal_avg_ipt | seasonal_avg_ipt | christmas_donation | scholarship_donation | non_seasonal_avg_expenditure | non_seasonal_expenditure | non_seasonal_frequency | seasonal_avg_expenditure | seasonal_expenditure | seasonal_frequency | seasonal_areas_dining | seasonal_areas_marina | seasonal_areas_golf | non_seasonal_areas_dining | non_seasonal_areas_marina | non_seasonal_areas_golf | yom | dist_km |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| non_seasonal_avg_ipt | 1 | | | | | | | | | | | | | | | | | |
| seasonal_avg_ipt | 0.43 | 1 | | | | | | | | | | | | | | | | |
| christmas_donation | -0.23 | -0.2 | 1 | | | | | | | | | | | | | | | |
| scholarship_donation | -0.17 | -0.11 | 0.49 | 1 | | | | | | | | | | | | | | |
| non_seasonal_avg_expenditure | -0.01 | 0.01 | -0.05 | -0.04 | 1 | | | | | | | | | | | | | |
| non_seasonal_expenditure | -0.22 | -0.18 | 0.11 | 0.08 | 0.57 | 1 | | | | | | | | | | | | |
| non_seasonal_frequency | -0.41 | -0.32 | 0.2 | 0.16 | -0.07 | 0.29 | 1 | | | | | | | | | | | |
| seasonal_avg_expenditure | 0.04 | 0.11 | -0.11 | -0.07 | 0.37 | 0.29 | -0.09 | 1 | | | | | | | | | | |
| seasonal_expenditure | -0.23 | -0.22 | 0.12 | 0.08 | 0.29 | 0.63 | 0.35 | 0.46 | 1 | | | | | | | | | |
| seasonal_frequency | -0.4 | -0.41 | 0.27 | 0.17 | 0.01 | 0.32 | 0.69 | -0.14 | 0.49 | 1 | | | | | | | | |
| seasonal_areas_dining | -0.29 | -0.53 | 0.3 | 0.17 | -0.19 | 0.03 | 0.24 | -0.26 | 0.08 | 0.35 | 1 | | | | | | | |
| seasonal_areas_marina | -0.21 | -0.17 | 0.08 | 0.05 | 0.31 | 0.43 | 0.13 | 0.14 | 0.33 | 0.32 | 0 | 1 | | | | | | |
| seasonal_areas_golf | -0.27 | -0.35 | 0.18 | 0.11 | -0.07 | 0.11 | 0.29 | -0.16 | 0.18 | 0.48 | 0.4 | 0.09 | 1 | | | | | |
| non_seasonal_areas_dining | -0.5 | -0.41 | 0.28 | 0.17 | -0.12 | 0.09 | 0.37 | -0.18 | 0.12 | 0.39 | 0.6 | 0.03 | 0.34 | 1 | | | | |
| non_seasonal_areas_marina | -0.24 | -0.16 | 0.11 | 0.08 | 0.36 | 0.47 | 0.16 | 0.09 | 0.29 | 0.33 | 0 | 0.73 | 0.08 | 0.04 | 1 | | | |
| non_seasonal_areas_golf | -0.26 | -0.26 | 0.13 | 0.08 | -0.12 | 0.04 | 0.38 | -0.13 | 0.14 | 0.43 | 0.22 | -0.01 | 0.48 | 0.29 | 0 | 1 | | |
| yom | -0.01 | -0.01 | 0.11 | 0.02 | -0.05 | 0 | 0.08 | -0.04 | 0.01 | 0.02 | 0.11 | -0.1 | -0.05 | 0.09 | -0.08 | 0 | 1 | |
| dist_km | 0.01 | -0.02 | 0.05 | 0 | 0.12 | 0.01 | -0.1 | 0.03 | 0.03 | -0.01 | -0.02 | 0.03 | 0 | -0.05 | 0.05 | -0.02 | 0 | 1 |

**Figure A1.** Correlation matrix of covariates.



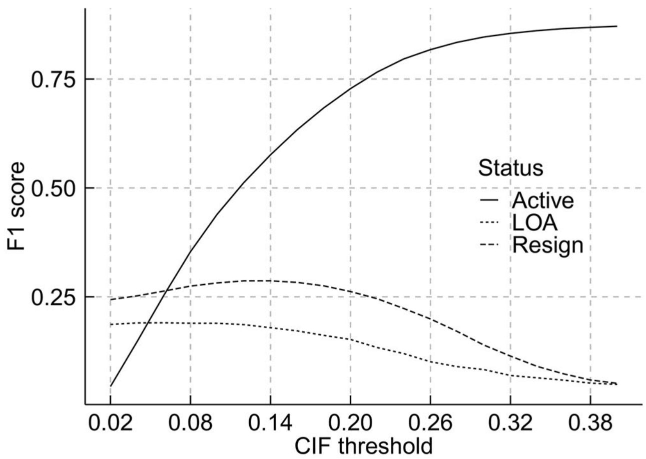




**Figure B1.** Precision and F1 score for different threshold CIFs for the event of active, LOA or resign.

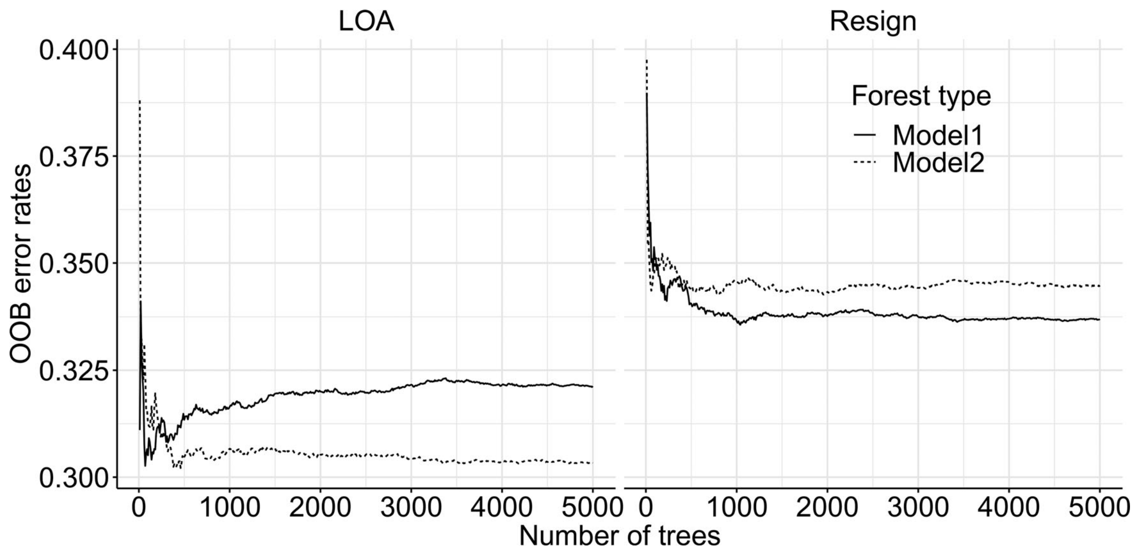



**Figure C1.** OOB error rates for final model (model1) and modified model (model2) for LOA (right) and resign (left) respectively. Hyper-parameters were consistent across both models.

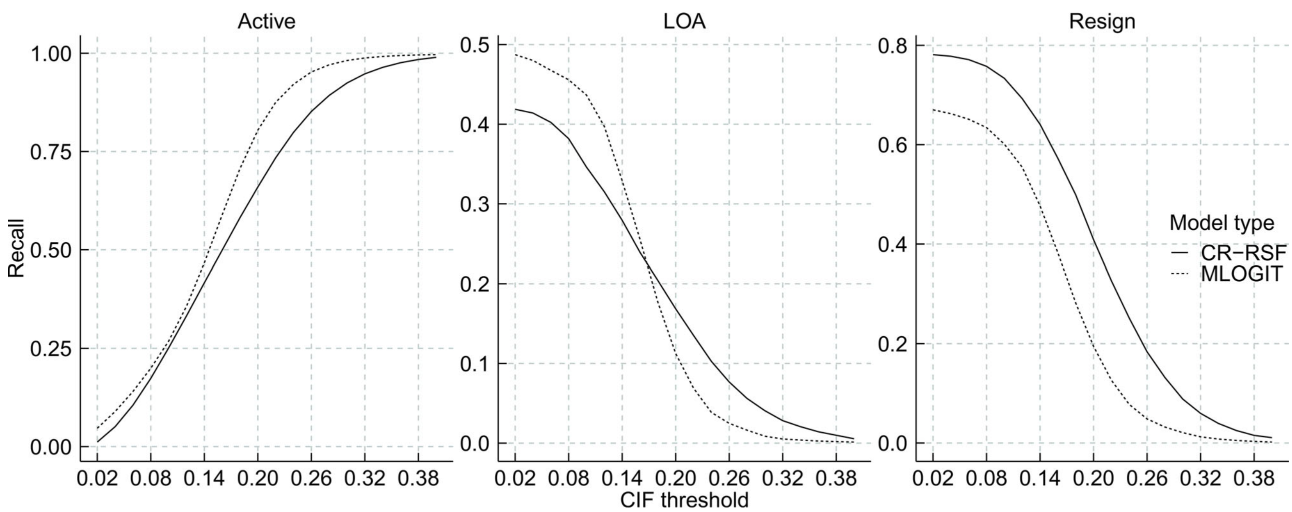



**Figure D1.** Recall for logit vs CR-RSF.

accuracy of active customers. Overall, precision and f1 score did not provide any additional information on accuracy, in comparison to recall.

## Appendix C error comparison of models with modified covariates

The covariates used in our final model for measuring churn could be specified differently. For instance, we could compute (i) the dollar values of expenditure for all cross-buying covariates instead of indicator functions, (ii) the dollar value of expenditure for donation variables instead of indicator functions, or (iii) exclude either total frequency, total expenditure or average expenditure if they were highly correlated. We constructed an alternative CR-RSF with these modified specifications. We computed the OOB error rates from this modified CR-RSF (model 2 in Figure C1) and the final CR-RSF model in the main manuscript (model 1 in Figure C1). As evident from the Figure C1, the OOB error rates are comparable for both models. Therefore, we did not include these modifications in the final model in our manuscript.

## Appendix D comparing accuracies of RSF and MLOGIT

Figure D1 demonstrates the recall, precision, and F1 score for CR-RSF and multinomial logit (MLogit). The results for recall, precision, and F1 score were computed based on 100 bootstrap cross validations, as outlined in Subsection 4.3. Since CR-RSF models for risk of churn across time and MLOGIT models for a single risk of churn, the estimates obtained from both models were not directly comparable. Therefore, we took the following steps to ensure the comparison of error statistics are consistent.

First, the CR-RSF provides estimates of churn for a customer for all time-periods (months) starting from 2009 and ending in 2016, for both events. However, the multinomial logit model provides a single churn prediction for both events, that is for the final month of 2016. Therefore, we only compare the MLogit prediction with the prediction obtained from CR-RSF that corresponds to the end of 2016.

Second, the Mlogit model provides the prediction of all three events: active, resign and leave. However, a survival model like CR-RSF provides the predictions of churning events only: resign and leave. Therefore, for a consistent comparison, we only use probabilities of resign and LOA to establish thresholds as outlined in Subsection 5.1 and assign customer status for the end of 2016 In other words, the predictions for both MLogit and CR-RSF are based on estimates of resign and LOA only.

Figure D1 shows that for predicting active customers, Mlogit outperforms CR-RSF in terms of recall scores across all thresholds. However, the results are opposite for LOA and resign. For higher values of threshold, CR-RSF is more accurate compared to Mlogit for predicting LOA. Moreover, CR-RSF outperforms Mlogit for all values of threshold for predicting resign. Similar results are observed for F1 score and precision (not shown). In summary, we observe that Mlogit outperforms CR-RSF for predicting active customers. However, for predicting customers who can churn, especially via multiple avenues, a survival based model such as CR-RSF, that uses time to churn, is better. Therefore, we only discuss survival-based models in the main manuscript, as the focus of this work is on churn prediction via multiple avenues.