Hazard rate models for core return modeling in auto parts remanufacturing

Akhilesh Kumar, Ratna Babu Chinnam, Alper Murat

Department of Industrial and Systems Engineering, Indian Institute of Technology Kharagpur, W. Bengal 721302, India

Department of Industrial & Systems Engineering, Wayne State University, 4815 Fourth Street, Detroit, MI 48202, USA

Article info

Article history:
Received 18 October 2015
Received in revised form 17 March 2016
Accepted 1 July 2016
Available online 5 July 2016

Keywords:
Remanufacturing
Core returns
Forecasting
Aftermarket
Hazard rate models

Abstract

Under growing consumer awareness and increasing legislation, firms are realizing the importance of including sustainability within their strategic objectives to promote their green image, enhance their corporate citizenship status, and also improve profit margins. Towards this end, sustainability through product remanufacturing is gaining momentum. However, a key complication for maintaining operational efficiencies during production planning and control of remanufacturing lies in the inability to accurately forecast core returns. These difficulties are mostly attributable to limited visibility and higher levels of uncertainty in reverse logistics. Despite significant advances in the remanufacturing literature over the last two decades, there is not yet a practical approach for modeling core return delay durations when the company is engaged in business with a large remanufacturing product catalog and many customer facilities. This is particularly true for suppliers that engage in both original equipment (OE) service as well as independent after-market (IAM) businesses. This research aims to address these limitations for suppliers by developing a range of hazard rate models for core returns duration modeling. Models are also validated using data from a large global automotive supplier.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

In today's global economy, firms are seeking any and every possible opportunity to differentiate themselves from competitors, to reduce their costs, and to add value to their supply chains and end customers. One increasingly popular option, under growing consumer awareness and increasing legislation, is to reintegrate used or returned product into the supply chain to regain the materials for economic and sustainability purposes (Schultmann et al., 2006). An important class of such “reverse” goods flows has to do with remanufacturing, which refers to activities that restore used products (known as “cores”) or their major modules to operational condition for use in place of new product or for use in other channels (e.g., spare parts). Remanufacturing has traditionally been prevalent in such industries as the automotive, electrical equipment, furniture, machinery, tires, and toner cartridges. For nearly two decades, the U.S. Environmental Protection Agency (EPA) has been advocating the practice of remanufacturing as economical, energy-efficient and environmentally friendly approach to reduce industrial waste (U.S. EPA, 1997). According to the Remanufacturing Institute, manufacturers of refurbished and remanufactured products save on average 85% of energy use, 86% of water use, and 85% of material use as compared to manufacturing a new good.1

Another important reason for improving reverse logistics is to cope with returns that have become endemic in many industries. For example, according to a Consumer Electronics Industry survey by the Reverse Logistics Executive Council, the average return rate is 8.46% in the high-tech industry (Thrikutam and Kumar, 2004), with return rates as high as 20% for certain product segments. While efficient management of these returns is necessary for obvious reasons, sustainability considerations also warrant the effective management of used product returns (e.g., end-of-use and end-of-life returns), often carried out by suppliers with the aid of original equipment manufacturers (OEMs). In comparison with forward supply chain operations, reverse logistics associated with remanufacturing is often far more complex. It not only deals with multiple types of returns from different sources that bring about various uncertainties (e.g., timing/location of return, volume of

---

1 http://www.reman.org
returns, quality of returns), but also has to address the complexities associated with remanufacturing production planning and control (Guide, 2000).

In the automotive industry (and several other industries), production parts can be broadly divided into Original Equipment (OE) parts and Aftermarket parts. OE parts refer to parts used in producing new vehicles, whereas, aftermarket parts refers to those traded after original equipment sale, which includes both OE service (for vehicles under OEM warranty) and independent aftermarket (IAM) services. The automotive aftermarket industry is estimated at $198B annually in the U.S. alone with IAM sales estimated at $142B. While the automotive remanufacturing business was traditionally dominated by IAM companies, high profit margins and growing pressures to improve corporate citizenship are encouraging more OEM and Tier-1 suppliers to pursue remanufacturing. According to a survey (Inmar, 2009), end-of-use 3

encouraging more OEM and Tier-1 suppliers to pursue remanufacturing. According to a survey (Inmar, 2009), end-of-use core return rates in the automotive industry are known to vary between 5% and 25%. Given the returns and the size of the aftermarket business, there are tremendous opportunities to engage in remanufacturing. While the automotive remanufacturing business was traditionally dominated by IAM companies, high profit margins and growing pressures to improve corporate citizenship are encouraging more OEM and Tier-1 suppliers to pursue remanufacturing. However, a key complication in maintaining efficiencies is the difficulty in production planning and control of remanufacturing activities.

Our setting is as follows. A supplier receives orders for remanufacturing parts from the “customer” which can be the aftermarket parts retailers/distributors (e.g., NAPA stores), OE service (e.g., orders from dealer service centers), and parts operations organizations (e.g., GM Service Parts Organization). In shipping the order (i.e., a remanufactured unit), the supplier often imposes a “core charge” on the customer, a debit that will be credited to the customer upon receiving the defective core. The supplier issues a return material authorization (RMA) in shipping the order to facilitate return shipment of cores. Efficient production and inventory management of remanufacturing parts for the supplier heavily impinges on the ability to accurately model core return durations from customers (besides forecasting demand for remanufacturing parts and securing cores from the open market, as necessary). There are several challenges facing these suppliers, including, the volume and diversity of customers, differences among individual customer warehouses in returning cores, large remanufacturing product catalog (product variety), changing customer behavior (often improving core return delays), and data sparsity.

Overall, while the extant literature offers a few returns forecasting models (Krapp et al., 2013; Ilijin and Gupta 2010), they are not practical for many suppliers, such as our collaborator Delphi Product & Service Solutions that provides replacement parts and services to the automotive aftermarket. There are several reasons for this: 1) they tend to model residence time distribution (in our setting it is equivalent to forecasting demand for remanufacturing parts), 2) are mostly applicable for forecasting returns of individual products/SKUs whereas these companies have extensive remanufactured product catalogs, and 3) make strong assumption that either core return probability over any period is known in advance or functional form of the return delay duration is known. Furthermore, the historical data in our case is simply too sparse to facilitate modeling and calibration of models for individual SKUs.

Thus, it is imperative to develop effective and scalable methods for modeling core return durations for these kinds of suppliers. A key characteristic of the core return datasets is right-censoring, meaning that at any given time, only a fraction of the returns are observed with the rest outstanding (not yet returned). In the absence of rich/reliable historical datasets, effective modeling of return duration times under censoring warrants survival analysis techniques. Conventional regression modeling approaches (e.g., duration time regression and logistic regression) are known to have several short-coming: 1) use of duration time regression in the face of censoring may lead to biased estimates of the covariate effects and 2) are inappropriate when there are time varying covariates. Furthermore, it is often the case that the effects of model “covariates” are non-stationary (e.g., return durations often improve over the product life-cycle and experience with remanufactured product). In examining duration time data, the three common objectives are (Helsen & Schmittlen 1993): to study the effects of covariates, learn the dynamics of duration, and duration time forecasting. The literature suggests that hazard rate regression models can overcome the above listed shortcomings while achieving all three objectives within a single tractable class of duration time models (Gupta, 1991; Jain and Villacissim, 1991; Helsen and Schmittlen, 1993; Hosmer et al., 2008). Further, Helsen and Schmittlen (1993) established that hazard rate regression models outperform conventional regression models in terms of stability of the estimates, face validity of parameter estimates, and predictive accuracy. Another advantage of hazard rate models, over regression models, is that they provide predictions that are probabilistic or expectation based. In the context of returns forecasting, the supplier might be interested in the expected number of returns by a certain date (or within a date range), which can be estimated based on the hazard rate distribution of outstanding RMAs. Further, continuous assessment of this prediction supports operational tasks such as inventory management and strategic decisions such as timing the launch of remanufactured products.

In particular, we report the evidence for the effectiveness of hazard rate models to estimate return delay distribution in the context of remanufacturing for suppliers. We studied various types of hazard rate modeling techniques and their appropriateness. Further, we discuss solutions to settings when the underlying assumptions are suspect, when covariates have time-varying effects, or there is randomness in the effects of the covariates. To the best of our knowledge, this is the first article to explore all these issues in the context of returns duration modeling for remanufacturing.

The rest of the paper is organized as follows: Relevant literature is discussed in Section 2. Proposed modeling approach is discussed in Section 3. A real world case study is presented in Section 4. Results and discussions are presented in Section 5. Finally, we conclude and provide directions for future research in Section 6.

2. Literature

Over the last two decades, there has been significant research in the area of remanufacturing and reverse logistics. Guide (2000) carried out an extensive survey of the remanufacturing literature and identified future research needs. Based on extant literature, he divided remanufacturing research into five broad categories: forecasting, reverse logistics, production planning and control, inventory control and management, and general. In recent years, the last four complications are being addressed extensively (e.g., Aras and Aksen, 2008; Barba-Gutierrez et al., 2008; Inderfurth, 2004; Krikke et al., 2008; Li et al., 2009; Takahashi et al., 2007; Tang and Grubbström, 2005; Wang et al., 2007). Since our research focus is on the first complicating characteristic, modeling return durations for facilitating effective forecasting, we encourage


3 Delphi Product & Service Solutions is a sub-division of Delphi Corporation, a leading global automotive supplier (http://delphi.com)
readers to refer a survey by Ilgin and Gupta (2010) for research in the other categories.

Toktay et al. (2000) discussed the role of forecasting in managing product returns and argued how predicting returns influences decisions at the strategic, tactical, and operational levels. They also indicated that there are few documented examples dealing with core returns duration modeling and forecasting in reverse logistics. Ilgin and Gupta (2010) and Krapp et al. (2013) reiterated this statement by citing a very few notable publications. In addition, most of these publications assumed that the core return probability over any period is known in advance (e.g., Goh and Varaprasad, 1986; Kelle and Silver, 1989). On a positive note, much of the recent literature is exploiting the fact that future end-of-use/end-of-life returns are a function of past sales. Goh and Varaprasad (1986) are credited for being the first to develop such a model. They proposed a transfer function model to estimate beverage product bottle returns using Box and Jenkins (1976) time-series techniques to compute life-cycle parameters. Kelle and Silver (1989) proposed four forecasting techniques based on available “information sets” to estimate the “net demand” during production planning and control for the lead time of reusable containers. Toktay et al. (2000) considered a queuing network based approach to achieve an optimal ordering policy for Kodak’s single use-camera. They utilized a Bayesian estimation and expectation optimization approach to forecast returns in trackable as well as untrackable cases. Although their method does not require known return rate, it makes an assumption regarding the shape of the lag distribution as geometric and negative binomial distribution. Krapp et al. (2013) emphasized on the importance of the dependencies between sold quantities and the resulting future returns which may lead to increased total average costs. They applied a Bayesian estimation technique similar to Toktay et al. (2000).

The effects of the covariates are captured through $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$, which correspond to the effect of covariate $x_i$ on the hazard rate, and $h_0(t)$ is the baseline hazard function. The multiplicative term $\varphi(x_i, \beta)$ adjusts the baseline hazard up or down proportionately and thus the estimation of $\beta$ is referred to as the proportional hazards regression (PHR). Commonly the

3. Hazard rate models for core return duration modeling

As stated earlier, core returns can be modeled using duration time modeling within a single tractable class — hazard rate models (also known as hazard function models). In this section, we describe the use of proportional hazard rate models for modeling core return durations. See Hosmer et al. (2008) for a comprehensive review of survival analysis methods and regression modeling of time-to-event data, including hazard rate models.

To accurately assess the effect of covariates (e.g., distinct customers and their facilities, product type, product size, and product features) and time-dependency in core return durations, it is important to control for the impact of sample selection (censoring) biases and the non-stationary nature of covariates. Hazard rate models provide an effective mechanism to tackle these issues.

3.1. Proportional hazard rate model

3.1.1. Notations

$T_i$: Event time

$f(t)$: Probability distribution function

$F(t)$: Cumulative distribution function

$x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})$: Covariate set

$h(t)$: Hazard function

$S(t)$: Survival function

$h(t|x_i)$: Hazard rate at time $t$ for an individual having covariate $(x_{i1}, x_{i2}, \ldots, x_{ik})$

$\beta = (\beta_1, \beta_2, \ldots, \beta_k)$: Estimated coefficients for covariates

$\mathcal{R}(t)$: Risk set

$L(\beta, \mathcal{R}(t))$: Partial likelihood

In hazard function based modeling, the times at which the event occurs (i.e., core returns) are assumed to be a realization of some random process. Accordingly, let $T_i$ denote the event time for the return of a particular core $i$, having a probability distribution p.d.f $F(t)$ and c.d.f. $F(t)$. In survival analysis, it is common to work with hazard function $h(t)$ and survival function $S(t)$ in place of $f(t)$ and $F(t)$, respectively. The survival function is $S(t) = 1 - F(t)$. The hazard rate function is $h(t) = f(t)/S(t)$ and measures the likelihood of the event happening at time $t$, given that it has not happened yet. The slope of the hazard function provides qualitative information on the duration time dynamics. In particular, if the likelihood of the event happening at time $t$ does not depend on $t$ (e.g., event happens randomly in time), then slope of hazard function is constant. For instance, when the durations are exponentially distributed, $f(t) = \lambda e^{-\lambda t}$, the hazard rate $h(t) = \lambda$ is constant, explaining the memory-less property of the exponential distribution.

The proportional hazard rate regression models developed originally by Cox (1972) enable incorporating covariates in duration time models. Denote $h(x_i)$ as the hazard rate at time $t$ for an individual having covariate values $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})$ at time $t$. The proportional hazard regression model is composed of two components representing the time-dependency of duration time $h(t)$ and effect of covariates $\varphi(x_i, \beta)$. The hazard rate function is then expressed in the multiplicative form as follows:

$$h(x_i) = h_0(t)\varphi(x_i, \beta)$$

(1)

The effects of the covariates are captured through $\beta = (\beta_1, \beta_2, \ldots, \beta_k)$, which correspond to the effect of covariate $x_i$ on the hazard rate, and $h_0(t)$ is the baseline hazard function. The multiplicative term $\varphi(x_i, \beta)$ adjusts the baseline hazard up or down proportionately and thus the estimation of $\beta$ is referred to as the proportional hazards regression (PHR). Commonly the
covariates’ effects term is modeled as an exponential function \( h(t; \beta) = h_0(t) \exp(\beta \mathbf{x}) \). Note that with the exponential function, the 
\[ \exp(\beta \mathbf{x}) \geq 0 \] is ensured.

### 3.2. Proportional hazard regression

There are two common approaches to estimating the hazard model: 1) semi-parametric and 2) parametric estimation. In the semi-parametric approach, we assume that \( h_0(t) \) can take any arbitrary shape and the estimation is performed through “partial likelihood” maximization (Cox 1972). In contrast, the parametric approach assumes a functional form for the baseline hazard \( h_0(t) \) (e.g., the Weibull distribution).

#### 3.2.1. Semi-parametric modeling (cox proportional hazard rate model)

The Cox proportional hazard (PH) model is one of the most widely used tools in survival analysis due to its efficiency and flexibility. This popularity is attributable to the semi-parametric nature of the model. Cox’s major contribution was to suggest an estimation technique (partial likelihood maximization) to estimate regression coefficients \( \beta \) while allowing for a general hazard function (Cox 1972). He also suggested that this can only result in a slight loss of information about \( \beta \). In fact, it is evidenced that maximizing the partial likelihood results in very efficient estimates of \( \beta \) (Efron, 1977; Oakes, 1977) and the partial maximum-likelihood estimation (MLE) is consistent and asymptotically normal under general conditions (Tsiatis, 1981).

At a given event time \( T_i = t \) of individual \( i \), the partial likelihood considers the relative risk of individual \( i \) realizing the event. Other individuals who had not yet experienced the duration event until time \( t \) are considered being at risk and forms the “risk set” \( \Omega(t) \). Note that individual \( i \) is also in the risk set. Hence, the partial likelihood is the probability that individual \( i \) is the one that experiences the event with duration \( T_i = t \) among those who has \( T_j > t \). To calculate this partial likelihood, the hazard rate \( h(t) \) can be used. The hazard rate measures each individual’s risk of realizing the event at time \( t \) provided that it has not happened until then. Let \( h(t) \) denote the hazard rate of individual \( i \). Then the relative risk (i.e., partial likelihood) can be calculated as

\[
L(t, i \in \Omega(t)) = \frac{h(t)}{\sum_{j \in \Omega(t)} h(t)} \tag{2}
\]

Using the proportional hazards model and canceling the common baseline hazard rate \( h_0(t) \),

\[
L(t, i \in \Omega(t)) = \frac{\exp(\beta \mathbf{x}(i))}{\sum_{j \in \Omega(t)} \exp(\beta \mathbf{x}(j))} \tag{3}
\]

where \( \mathbf{x}(i) \) is the covariate vector for individual \( i \) at time \( t \). The estimate for \( \beta \) and individual \( i \) is obtained by maximizing the product of expression in (3) over all observed durations including the right-censored observations. The Cox-model is commonly used because it provides flexibility, works well in practice, and is reasonably robust to modest departure from the proportional hazard assumption. Further, in many cases, variables can be transformed to show approximate proportional hazard (discussed in Section 5.3).

#### 3.2.2. Parametric modeling

Parametric models assume that the base-line hazard function \( h_0(t) \) follows a known functional form (e.g., Exponential, Weibull, Logics). All parametric models can be fit by maximizing the appropriate likelihood function. Given \( T_i \) (the event durations or censoring time) and \( R_i \) indicating event occurrence (\( R_i = 1 \)) and censored cases (\( R_i = 0 \)), the following likelihood function is maximized under non-informative censoring:

\[
L = \prod_i h(t = T_i | r(i))^{R_i} S(t = T_i | r(i)) \tag{4}
\]

Computationally, the biggest advantage of a parametric model is that one can use full maximum likelihood to estimate the parameters, which, in turn, provides meaningful estimate of effects. Parametric models are a better choice if the modeler has good knowledge of the return process. The literature cautions the use of parametric models since prior knowledge is not always available. Nonetheless, this does not rule out the option of comparing parametric models against semi-parametric models.

### 4. Automotive case study

Our motivation for this case study originates from a Tier-1 auto parts supplier’s approaching us with their need to reliably and accurately forecast their core returns. As widespread in the automotive industry, this supplier has been relying on the average number of returns per year to plan for their operations and investment. Their needs and problem setting provided an ideal testbed for our research’s case study because of the multiplicity of their core SKUs as well as absence of any advanced knowledge of core return probability or the functional form of the return delay duration. Both of these reasons are common to most remanufactured parts/products in the automotive industry.

To establish the empirical performance of the proposed hazard rate models for modeling core return durations, we tested it on actual returns data for an Electronic Control Module (ECM). ECMS are critical subsystems in today’s modern vehicle designs and consist of CPUs and assorted signal inputs and outputs dedicated to controlling a component within the vehicle. There are different kinds of ECMS, which vary in complexity such as an Engine Control Unit managing the power-train system efficiency, an Anti-lock Braking (ABS) Control Unit monitoring the vehicle speed and brake fluid, and a simple body module controlling the automatic door locks or power windows.

#### 4.1. Dataset

The ECM core returns data was collected from our collaborator Delphi Product & Service Solutions over a period of several years. The dataset covered core return transactions for hundreds of ECM part numbers and more than 30 distinct customers.\(^4\) The dataset provided did not contain any information regarding product attributes (e.g., ECM type, size, features, price) to facilitate more effective parameterization of the hazard rate model. The same was true with respect to customers and their facilities. We only had access to product ID, customer ID, product shipment dates, and core return dates, limiting the ability of the model to support data sparsity.

Preliminary data cleaning revealed that there were many customers who never returned cores. Also, in some of the cases, the customers did business with Delphi for a very small time period. Thus, in the remaining analysis, we only considered customers who returned at least 10 core units. Fig. 1 shows history of core return delays for a particular Delphi ECM product family. The axes scales are hidden throughout this document for reasons of confidentiality.

\(^4\) Due to confidentiality reasons, further delineation of the dataset is restricted by the case study’s Tier-1 supplier.
4.2. Nomenclature

This section provides necessary nomenclature to facilitate core return duration modeling under right censoring in an IAM setting.

\[ D_a \] Census Date
\[ D_s \] Shipment Date
\[ D_k \] Day product was returned (1 if it is returned, 0 otherwise)
\[ R, R_i \] Indicator variable, 1 if product is returned, 0 otherwise
\[ T, T_i \] Duration time, from \( D_a \) until \( D_k \) or \( D_s \), depending on whether core is returned or not
\[ P, P_i \] Product Index, \( i = 1, 2, 3 \ldots \)
\[ C, C_i \] Customer Index, \( i = 1, 2, 3 \ldots \)

Now, we can construct a hazard rate model with a set of covariates \( X = \{ P, C \} \), and dependent variable \( T \) with censoring \( R \) as follows:

\[ h(T, R|X) = h_0(t) e^{\beta P + \beta C} \]  (5)

5. Results and discussion

5.1. Parameter estimates

To assess the suitability of the proposed hazard rate models, we chose the popular semi-parametric Cox proportional hazard model as well as the parametric Weibull hazard rate model, which allows for both increasing as well as decreasing hazard rate over time. Firstly, we studied the effect of covariates and their significance. Our initial analysis revealed that none of the products were statistically significantly different from each other in their return patterns. Thus, we assumed that all ECM products are identical and chose customer index as the only covariate for return duration modeling.

Table 1 summarizes the model results obtained for both the parametric and semi-parametric (Cox PH) hazard rate models. Fig. 2 reports the resulting core return survival probabilities with respect to time (i.e., probability of not receiving a core from the customer) as estimated by the Cox PH model for the different customers. It is apparent in Fig. 2 that most of the customers are rather different in their core return duration distributions, with Customer 6 being the fastest in returning the cores.

Table 1. Estimates of model coefficients for ECM returns data.

<table>
<thead>
<tr>
<th>Weibull Model</th>
<th>Cox PH Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Intercept</td>
<td>(-6.12)</td>
</tr>
<tr>
<td>Customer 2</td>
<td>(-0.50)</td>
</tr>
<tr>
<td>Customer 3</td>
<td>(-0.40)</td>
</tr>
<tr>
<td>Customer 4</td>
<td>(-0.10)</td>
</tr>
<tr>
<td>Customer 5</td>
<td>(-0.54)</td>
</tr>
<tr>
<td>Customer 6</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Customer 7</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Log(Scale)</td>
<td>(-0.28)</td>
</tr>
</tbody>
</table>

Log-likelihood = \(-6298.9\) \( R^2 = 0.049 \)
is not similar to that of in the linear models. For instance, as Hosmer et al. (2008) states that a perfectly adequate model may provide a very low $R^2$ due to a high percent of censored data. In our results, we present the $R^2$ values for the purpose of showing that frailty model is performing better than Cox-PH model.

5.2. Validation

In this section, we present various performance measures to establish the validity of the two models. Readers should note here that our intention is to present the validity of these two models within the class of proportional hazard rate models but not to promote the use of one model over other. This is because direct comparisons of these models are not impartial, since: 1) the parametric model is based on event times whereas the Cox PH model is based on rank of event times, 2) and scales of the parameters may differ significantly. Thus, all the comparative studies presented in this section intend to show the suitability and the relative performance of the models. We divide our validation process into two parts: 1) stability and efficiency of the estimates, and 2) predictive performance of the models.

5.2.1. Stability and efficiency of the estimates

To check the stability of the estimates obtained using the proposed models, we performed experiments by sub-sampling data from the entire ECM dataset. To pursue this analysis, we have referred to the methods proposed by Helsen and Schmittlen (1993) and considered three sample datasets: 1) complete data as dataset (I), 2) 50% of complete dataset as dataset (II), and 3) the remainder 50% of complete dataset as dataset (III). One should note here that taking a totally random sample may compromise class balance, since we have returns as well as non-returns. To retain uniformity across samples, we considered 50% of returns and 50% of non-returns in both dataset II and dataset III. We re-estimated all the models for the three datasets. To evaluate the relative efficiency of the estimates, we evaluated the standardized measures of variability, $SV = \sigma / \hat{\mu}$, for all the models in all the samples. $SV$ is analogous to the coefficient of variation, where cases with parameter estimates close to zero are emphasized (Nardi and Schemperi, 2003).

Table 2 presents the estimated coefficients and SVs (in parentheses) for each sample for both models. Table 2 shows favorable performance of the models in term of stability of the parameter estimates. There is remarkable consistency (aside from minor discrepancies) between the model results of datasets II and III. More interestingly, there is very small departure from estimates obtained from complete dataset versus when sample size is reduced to 50%. Also, there is not a single discrepancy in the sign of coefficients. This indicates the suitability of proportional hazard rate models for modeling returns from the standpoint of stability of the parameter estimates, even when sample sizes are reduced to 50% of the original sample.

To compare relative efficiency of the models, we compared SVs of parameter estimates. One can easily see that in most of the cases (4 out of 6 for all three samples), Cox PH model’s SVs were closer to zero in contrast to those of the Weibull model. Thus, we can conclude that the Cox PH model performed better than the Weibull model based on standard measures of variability.

5.2.2. Predictive performance of the models

We compared each model’s prediction vis-à-vis the observed returns. In order to facilitate these comparisons, we considered two performance measures: hit rates (Helsen and Schmittlen, 1993) and mean square errors (MSE) in forecast.

Hit rates can be defined as percentage of returns correctly classified. To calculate hit rates, we need the hazard rate model forecasts for median duration and the observed returns within the median duration. Hazard rate model forecasts for median duration imply computation of a time point at which the survival probability crosses the threshold of 0.5. It is not always possible that an observation coincides with a 0.5 value on survival function, thus, when required, we performed a linear interpolation to produce forecast for the median duration. Table 3 presents hit rates for the two models for all the samples. It should be noted that, for each sample, we used a model calibration and validation ratio of 70% to 30%. For Cox PH model, the hit rates are as high as 90%. The Cox PH model performed remarkably well compared to Weibull model. This difference can be better understood by analyzing the baseline survival plots for both models (Fig. 3). The Cox PH model provides a better fit to the data due to its flexibility in contrast to the more rigid Weibull model.

Next, we consider the mean square error of the return duration forecast. To achieve this objective, we partitioned the full dataset into five datasets based on five consecutive time periods. The same procedure, as used to forecast the mean duration, is then used to generate the forecasts for each of the time periods. In this scenario, we compare the number of estimated returns with observed returns for each time period. Table 4 presents the overall MSE of forecast for each model for each customer. The Cox PH model once again out-performed the Weibull model. In light of these results, we conclude that Cox model performed better than Weibull for this particular case study. However, the results for different applications will be strongly dependent on the size of the datasets. For overly small datasets, the parametric models tend to fare better given the need to estimate few baseline hazard function parameters.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Weibull Model</th>
<th>Cox PH Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 2</td>
<td>-0.50 (0.45)</td>
<td>-0.47 (0.67)</td>
</tr>
<tr>
<td>Customer 3</td>
<td>-0.40 (0.38)</td>
<td>-0.46 (0.47)</td>
</tr>
<tr>
<td>Customer 4</td>
<td>-0.10 (0.67)</td>
<td>-0.13 (0.69)</td>
</tr>
<tr>
<td>Customer 5</td>
<td>-0.53 (0.25)</td>
<td>-0.55 (0.35)</td>
</tr>
<tr>
<td>Customer 6</td>
<td>0.23 (0.31)</td>
<td>0.23 (0.43)</td>
</tr>
<tr>
<td>Customer 7</td>
<td>0.31 (0.34)</td>
<td>0.32 (0.45)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Weibull Model</th>
<th>Cox PH Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: Complete dataset</td>
<td>67.6%</td>
<td>91.3%</td>
</tr>
<tr>
<td>II: Calibration-dataset</td>
<td>66.3%</td>
<td>91.1%</td>
</tr>
<tr>
<td>III: Validation-dataset</td>
<td>68.3%</td>
<td>91.5%</td>
</tr>
</tbody>
</table>
For each time-period. For the $\beta$ can take functional forms such as
Section 5.2.2. Fig. 4 plots the effective
consider Customer 5 and different time-periods described in
not easily interpretable. To better understand the dynamics, let us
tical tests are found to be signiﬁcant, the parameter estimates are
time-by-covariate interactions into the model. Although all the statis-
terms. Time-by-covariate interactions can be captured by a simple
the Cox PH model by including time-by-covariate interaction
the proportional hazard assumption is suspect, we re-estimated
though there was modest departure from proportionality. Since,
residuals test. Test results revealed that there is evidence against
To check time trend in the covariates, we performed a Schoenfeld
predictive performance? According to Cox and Oakes (1984), when
some customers but, for others, Weibull is better in terms of
baseline hazard rate in the absence of other
covariates (Table 6). Results obtained reveal that there is evidence against
for different time periods (Customer 5) under Cox PH with time-
by-covariate interactions.

Table 6
Model coefficient estimates for Cox PH with random effects (Frailty Model).

<table>
<thead>
<tr>
<th>Customer</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer 2</td>
<td>-0.83</td>
<td>0.30</td>
<td>0.01</td>
</tr>
<tr>
<td>Customer 3</td>
<td>-0.68</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td>Customer 4</td>
<td>-0.15</td>
<td>0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>Customer 5</td>
<td>-0.52</td>
<td>0.18</td>
<td>0.00</td>
</tr>
<tr>
<td>Customer 6</td>
<td>0.28</td>
<td>0.10</td>
<td>0.01</td>
</tr>
<tr>
<td>Customer 7</td>
<td>-0.21</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>Frailty (Product)</td>
<td>0.198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance of Random Effect</td>
<td>0.198</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.151</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Another possible explanation for the deviation from proportionality of the hazard could be due to some random effect caused by other covariates which are excluded. This can be better captured by random-effect models. In Cox PH regression, a frailty model is a random effects model for time variables, where the random effect (the frailty) has a multiplicative effect on the hazard (Hougaard, 1995). To examine the presence of this phenomenon, we performed Cox PH regression with customer and ‘frailty due to products’ as covariates (Table 6). Results obtained reveal that ‘frailty due to products’ is highly significant with a variance of 0.198. Although our initial experiments suggested that product type by itself was statistically insignificant, its random effect is found to be significant. Further, $R^2$ value increases from 0.049 to 0.151.
5.4. Discussion and managerial insights

Our analysis reveals that customers are delaying the core returns and in some instances almost indefinitely, instead of returning the core right after receiving a remanufactured unit shipment from the supplier. A possible reason is that customer facilities are ordering a remanufactured unit for storage rather than use it for an immediate customer. Thus, they cannot return the core until they receive a customer vehicle needing an ECM replacement. Other reasons could be the presence of independent core-collectors buying the cores from repair shops at a higher price than the core charge imposed by the supplier. We also suggest that results obtained by time-by-covariate interaction models can help in understanding the return behavior of each customer over a particular time horizon. Thus, from managerial perspective, it is very important to devise an attractive incentive mechanism such that this delay can be reduced.

Based on frailty model, we recommend that various product attributes, such as product size, weight, core deposit charge, and demand should be incorporated in the model. This could lead to improved model fidelity and explanation power and also help improve the scalability of the model for large product catalogs under data sparsity.

Overall, the time-by-covariate interaction model was able to explain the dynamics better than other models. However, from computational complexity point of view, other models were found to be far more superior. This is because, with time-by-covariate interactions, there is a baseline hazard rate for every time stamp. Thus, choosing one of these models requires a trade-off analysis between the computational complexity and the degree of accuracy desired. In our case, Cox PH model performed satisfactorily barring the modest departure from proportionality in meeting the requirements.

6. Conclusions

This paper presents a collection of models for modeling core return delay durations for aftermarket suppliers of re-manufactured product. This modeling approach is critical for effective forecasting of core returns and, in turn, for production planning and control of remanufacturing. The proposed modeling approach recommends hazard rate models for their efficacy and flexibility to support product attributes and customer features as covariates to improve the scalability of the models to support suppliers with extensive remanufactured product catalogs and diverse customers.

In particular, two proportional hazard rate models (one parametric and the other non-parametric) have been discussed. The models have been tested using a real dataset from a leading global automotive supplier with extensive remanufactured product lines. Our findings indicate that, Cox PH models are more effective than parametric Weibull models in terms of parameter estimates and predictive accuracy. However, the ultimate choice of the approach depends on the size of available datasets, experts/analysts knowledge about the return process, and other factors. While, we considered automotive parts returns, the presented models are capable of achieving higher level of scalability and can be employed for returns duration modeling in other industries. Future research can also investigate other hazard rate models besides the proportional hazard models discussed in this manuscript.

References

Clotey, T., Benton, W.C., 2014. Determining core acquisition quantities when products have long return lags. IIE Trans. 46 (9), 880–893.