# Analysis of Field Return Data With Failed-But-Not-Reported Events

Xin WANG[†], Zhi-Sheng YE[†], Yi-Li HONG[‡], and Loon-Ching TANG[†]

[†]Department of Industrial and Systems Engineering, National University of Singapore

[‡]Department of Statistics, Virginia Tech

**Abstract**

Warranty data contain valuable information on product field reliability and customer behaviors. Most previous studies on analysis of warranty data implicitly assume that all failures within the warranty period are reported and recorded. However, the failed-but-not-reported (FBNR) phenomenon is quite common for a product whose price is not very high. Ignorance of the FBNR phenomenon leads to an overestimate of product reliability based on field return data or an overestimate of warranty cost based on lab data or tracking data. Being an indicator of customer satisfaction, the FBNR proportion provides valuable managerial insights. In this study, statistical inference for the FBNR phenomenon as well as field lifetime distribution is described. We first propose a flexible FBNR function to model the time-dependent FBNR behavior. Then, a framework for data analysis is developed. In the framework, both semiparametric and parametric approaches are used to jointly analyze warranty claim data and supplementary tracking data from a follow-up of selected customers. The FBNR problem in the tracking data is minimal and thus the data can be used to effectively decouple the FBNR information from the warranty claim data. The proposed methods are illustrated with an example.

Key Words: Semiparametric estimation; Log-location-scale distribution; Maximum likelihood; Warranty; Field reliability; Reporting behavior.

# 1 Introduction

Warranty is an important part of a product. It affects purchase decisions of customers and profits of manufacturers since offering warranty causes additional costs (Huang et al., 2007). These costs include receiving, diagnosing, repairing, replacing, repackaging, restocking and/or reselling returned products. Generally, warranty costs account for 2% to 15% of the total sales price (Blischke et al., 2011). It is important to estimate the warranty cost accurately so that the right amount of money can be reserved to meet future warranty claims. The warranty costs depend on both product field reliability and customer reporting behaviors. A warranty service will be initiated only when (i) a product is sold, (ii) it fails within the warranty period, and (iii) its user claims the warranty. Therefore, the product reliability *per se*, the sales process, as well as the customer reporting behavior contributes to the overall warranty cost.

The primary reason to provide warranty is due to unreliability of the product. The probability of a field failure within the warranty period can be obtained based on probabilistic models for the failure pattern of the product as well as statistical methods for warranty data. Some studies proposed probabilistic models to depict the failure pattern by considering various factors such as different warranty policies (Chukova and Johnston, 2006), quality variations in manufacturing and assembly process (Nair et al., 2001), operating environment like temperature, humidity, vibration and pollution (Ye et al., 2013), use frequency (Hong and Meeker, 2010), multiple failure modes (Taylor and Peña, 2014), and recurrent events (Hu and Lawless, 1996). Alternatively, direct analysis of the warranty data gives a more straightforward estimate of the failure probability (Hsu et al., 2015). For a comprehensive overview of the warranty data analysis, see Wu (2013). The above work focuses more on the field reliability of a single unit. To forecast the number of warranty claims over time, the dynamic sales process should be taken into account. The homogeneous and non-homogeneous Poisson processes (Ja et al., 2002; Gurgur, 2011; Xie et al., 2014) are popular models for the sales over time.

In addition to the product reliability and the sales process, customer-reporting behaviors are also an important factor affecting the number of warranty claims. The behaviors include delay reporting (Kalbfleisch and Lawless, 1992; Rai and Singh, 2006), fraudulent reporting due to such reasons as misuse, and the failed-but-not-reported (FBNR) behavior (Wu, 2011). The FBNR behavior is a common phenomenon in warranty claims. It reduces the number of warranty claims and leads to an over-reservation of warranty money. Possible reasons for the FBNR phenomenon could be as follows. Due to rapid-technological innovation, products, especially electronic products such as cellphones, cameras, USB flash memories, and home entertainment systems, are easily obsolete. Some customers may shift their affection to new designs when their current products are still functioning, or the products can no longer meet the users' requirements. In these circumstances, a failure within warranty may not result in a warranty claim. Another possible reason is the difficulty in accessing warranty services. Instead of spending time waiting for claim handling and product repair, customers may simply purchase new ones when their products fail. In reality, the FBNR phenomenon could happen to most products. For instance, major mobile phone companies launch a new and improved generation device every six to twelve months. Impelled by better functions and models, customers tend to change cellphones more frequently (Wilhelm et al., 2011).

Most studies interpreted a product with no claim as with no failure. They implicitly assumed that the number of failures within the warranty period is equal to the number of warranty claims received. This assumption, however, may not hold in reality due to the FBNR phenomenon argued above. Generally, the number of warranty claims (observed failures) is smaller than that of actual failures. Warranty studies concerning the FBNR behavior are quite limited in the literature. Patankar and Mitra (1995) coined the behavior as the "partial redemption of warranty" or "partial warranty execution". They proposed two classes of warranty execution functions, which were essentially one minus the FBNR proportion, to model the time-dependent reporting behavior. Wu (2011) examined effects of the FBNR phenomenon on warranty reserves for a product having two failure modes. Xie and Liao (2013) derived the mean and variance of the total warranty and post-warranty repair demand given a constant FBNR proportion. However, the above studies mainly focus on probability modeling of the FBNR behavior. Statistical inference for the FBNR parameters in these probability models is not found. Obviously, the validity of assumptions on the FBNR behavior and the values of the FBNR parameters should be inferred from warranty data.

As the FBNR phenomenon indicates customer's dissatisfaction of the product or the warranty service, accurate estimation of the FBNR proportion not only helps manufacturers budget for the right warranty reserve, but also helps them better understand target users' sentiments. In this study, statistical inference on the FBNR behavior is investigated. The major source of data we can utilize is the warranty claim data from all customers. Nevertheless, the warranty claim data are quite messy. Besides the FBNR behavior, the data are often contaminated by many factors reviewed above, such as heterogeneous customer use behaviors. To analyze the warranty claim data, assumptions such as on the lifetime distribution and on the customer reporting behavior are needed. These assumptions are extremely difficult, if not impossible, to verify using the warranty claim data alone. Luckily, manufacturers often implement tracking studies where a number of customers are selected randomly and followed up closely. The major purpose is to learn customer experience and sentiment (Stevens, 2006, pp. 17). By-products of these studies are failure reports from these customers, called tracking data in our study. An example can be found in Hong and Meeker (2010). The tracking data provide additional information on product's field reliability. It may be reasonable to believe that there is no FBNR problem in the tracking data. Assumptions on the lifetime distribution and the FBNR behavior could be verified by jointly analyzing the tracking data and the warranty claim data. In this study, we first propose a flexible parametric FBNR function. Then a statistical inference framework for the FBNR phenomenon is developed through joint analysis of the warranty claim data and the tracking data.

The rest of the paper is organized as follows. Section 2 introduces the data of interest. A flexible model for the FBNR proportion is also proposed. Based on the setting, a statistical inference framework is proposed in Section 3. Section 4 develops a semiparametric approach that jointly analyzes the warranty claim data and the tracking data. Section 5 deals with parametric point and interval estimation. Section 6 discusses hypothesis test procedures for selecting an appropriate baseline lifetime distribution and a proper FBNR function. A simulation study is conducted in Section 7 to assess the performance of the proposed estimation methodology. Section 8 demonstrates the inference framework with a real example. Section 9 concludes the paper.

# 2 Problem Description

## 2.1 Available data

Two sets of data are used in our analysis. The first set is the warranty claim data for untracked units in the field, which is often well-maintained due to the development of modern technology on database management systems. Let $\tau$ be the warranty period. Let $X$ be the lifetime of a random unit and $C$ be the associated censoring time for the unit. The value of $C$ depends on $\tau$ as well as the sales date of the unit. Suppose that when the unit fails at age $t, t < C$, it will be reported to the manufacturer as a warranty claim with probability $1 - q(t), 0 \leq q(t) \leq 1$. This means that the FBNR probability $q(t)$ is a function of time. We will propose a flexible parametric form for $q(t)$ in Section 2.2 to capture the non-decreasing FBNR proportion overtime. Let $\delta$ be the failure-report indicator for the untracked unit. The indicator $\delta = 1$ if the unit fails before the censoring time $C$ and the failure is reported, and $\delta = 0$ otherwise. Furthermore, let $T$ be the observed service time for the unit, i.e., $T = \delta X + (1 - \delta)C$. The data observed from the untracked unit is $(T, \delta)$. Suppose that the warranty claim database contains $n$ untracked units in total. Therefore, there are $n$ realizations of $(T, \delta)$, denoted as $\mathbf{D}_{wc} = \{(t_i, \delta_i); i = 1, 2, \cdots, n\}$.

The second one is called the tracking data for a subset of sold units that are placed under a tracking study. By installing sensors in the units or following-up customers through email or phone, all failures within the observation period will be recorded and the FBNR problem is minimal. These failures are not counted into the warranty claim database. Note that the FBNR phenomenon could still happen when the tracked customers do not respond to the follow-up. Nevertheless, the proportion of unresponded customers is expected to be small and they are suggested to be excluded from the tracked group, which will further mitigate the FBNR problem within the remaining tracking population. It is assumed that the tracking sample is representative of the general population. Let $\tilde{\delta}$ be the failure-report indicator for a tracked unit with lifetime $X$ and censoring time $C$. Then $\tilde{\delta} = I_{\{X \leq C\}}$, where $I$ is the indicator function. This means that we assume no FBNR problem, and thus $\tilde{\delta}$ is equivalent to the censoring indicator. The observed data for the unit is $(T, \tilde{\delta})$. Suppose there are $m$ tracked units in total. Therefore, the observed tracking data are $\mathbf{D}_{tr} = \{(t_j, \tilde{\delta}_j); j = 1, 2, \cdots, m\}$.

## 2.2 Time dependent FBNR behavior

An appropriate form for the FBNR function $q(t)$ requires a good understanding of the reporting behavior. To model the time-dependent reporting behavior, Patankar and Mitra (1995) studied two classes of piecewise warranty execution functions, which are $1-q(t)$. As discussed in Patankar and Mitra (1995), a customer is more likely to execute a warranty claim if the product fails at an initial stage of the use. The above observation suggests that the FBNR proportion may be increasing over time. Moreover, as indicated in their model, the customer's willingness to report the failure could be relatively constant at the beginning of usage. Therefore, the FBNR proportion was assumed constant zero from time zero to a random change point, after which the FBNR proportion was assumed to follow a linear or an exponential function. Nevertheless, for low-price products such as thumb drives, the FBNR phenomenon could occur at the very beginning, i.e., $q(0) > 0$, which cannot be captured by their models. In addition, the piecewise functions they proposed are not smooth enough, which poses great difficulties in the statistical inference.

To circumvent the difficulty in estimation, we propose a flexible parametric form for the non-decreasing FBNR function as

$$q(t) = 1 - r_0 \exp\left[-(at)^b\right], \qquad t \geq 0, \tag{1}$$

where $r_0 \in [0, 1]$ is the initial report proportion at time zero, and $a \geq 0, b > 0$ are scale and shape parameters of the FBNR function.

The function (1) can have different shapes and is able to describe various reporting behaviors. For example, when the product life cycle is short, the FBNR function may

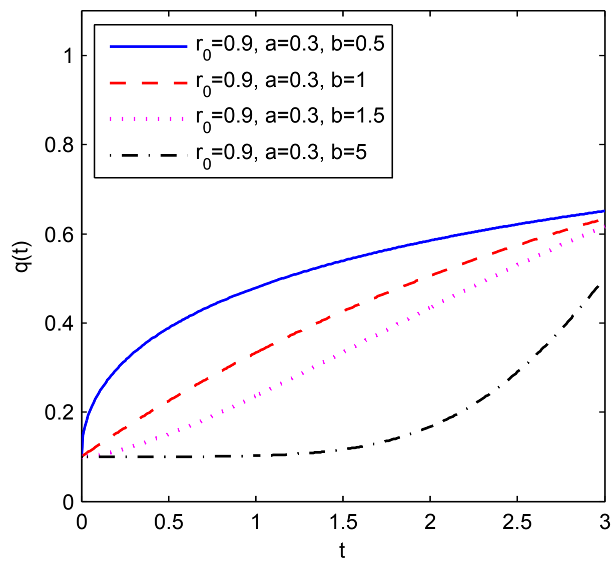


Figure 1: FBNR function $q(t)$.

increase slowly at the beginning and increase fast when the life cycle is approaching. If we let $b > 1$ in (1), then the function would give the desired property; see the dash-dot line in Figure 1. When the product is expensive or when it is easy to approach the manufacturer to claim the warranty, users are more likely to report failed items. In these cases, the FBNR problem may be rare at the beginning (i.e., $r_0 = 1$) and the FBNR proportion may be low during the whole warranty period. In some other cases, the FBNR proportion can be high. For nonrepairable products that are sold with pro-rata warranty policies, refund decreases with the failure time. Near the end of the warranty period, the refund is near zero, which may lead to a very high FBNR proportion. The FBNR function can also be relatively constant for a product that are with a long life cycle and are not expensive. In this case, we have $a = 0$. When $a = 0$ and $r_0 = 1$, we have $q(t) = 0$, meaning that all failures within the warranty period are reported. Therefore, the FBNR function is flexible enough to model a variety of FBNR problems. See Figure 1 for some examples.

# 3 A Framework for Data Analysis

In this section, a statistical analysis and inference framework for warranty data in the presence of the FBNR behavior is developed. Two key elements in the framework are the product lifetime distribution $F(\cdot)$ and the FBNR function $q(\cdot)$. After warranty claim data and tracking data are collected, the inference begins with semiparametric estimation that assumes a nonparametric lifetime model and a parametric FBNR function (1). See Section 4 for the detailed inference procedures. Results of the semiparametric inference are helpful in identifying an appropriate parametric lifetime distribution for the subsequent parametric inference. In addition, experience from lab tests, past designs and/or engineering knowledge is also useful in selecting the lifetime distribution. See Section 6.1 for detailed discussions about selection of a parametric lifetime model. The semiparametric inference also provides rough estimates of the three parameters of $q(\cdot)$ in (1). The estimates, together with the manufacturer's knowledge on the product, would be helpful in determining possible values of some parameters in $q(\cdot)$. For example, if the rough estimate of $r_0$ is close to 1 and the manufacturer believes no FBNR at the beginning, we may set $r_0 = 1$ in the subsequent parametric inference. Alternatively, quantitative methods such as hypothesis tests and the Akaike information criterion (AIC) can be used to test these possible values. See Section 6.2 for details of the tests. Based on the selected lifetime model and the FBNR function, both point and interval estimations for the model parameters could be readily done through joint analysis of the warranty claim data and the tracking data. Point estimators could be obtained by the maximum-likelihood (ML) method. See Section 5.2. Various methods such as the large-sample normal approximation and the bootstrap can be used to construct confidence intervals for parameters of interest. See Section 5.3.

The inference results provide managerial insights and can be used for decision support. For example, the lifetime distribution informs the reliability of the product in the field. As an indicator of customer's satisfaction of the product or the related warranty service, the FBNR function reflects the customer's sentiment on the product. If the product is still on sale at the time of analysis, then the results can be used to forecast warranty reserve for this product by considering both the product reliability and the FBNR behavior. If the product is no longer on sale, the results can be useful for products of next generation or similar products. A flowchart of the inference framework is given in Figure 2 to better visualize the inference procedures.

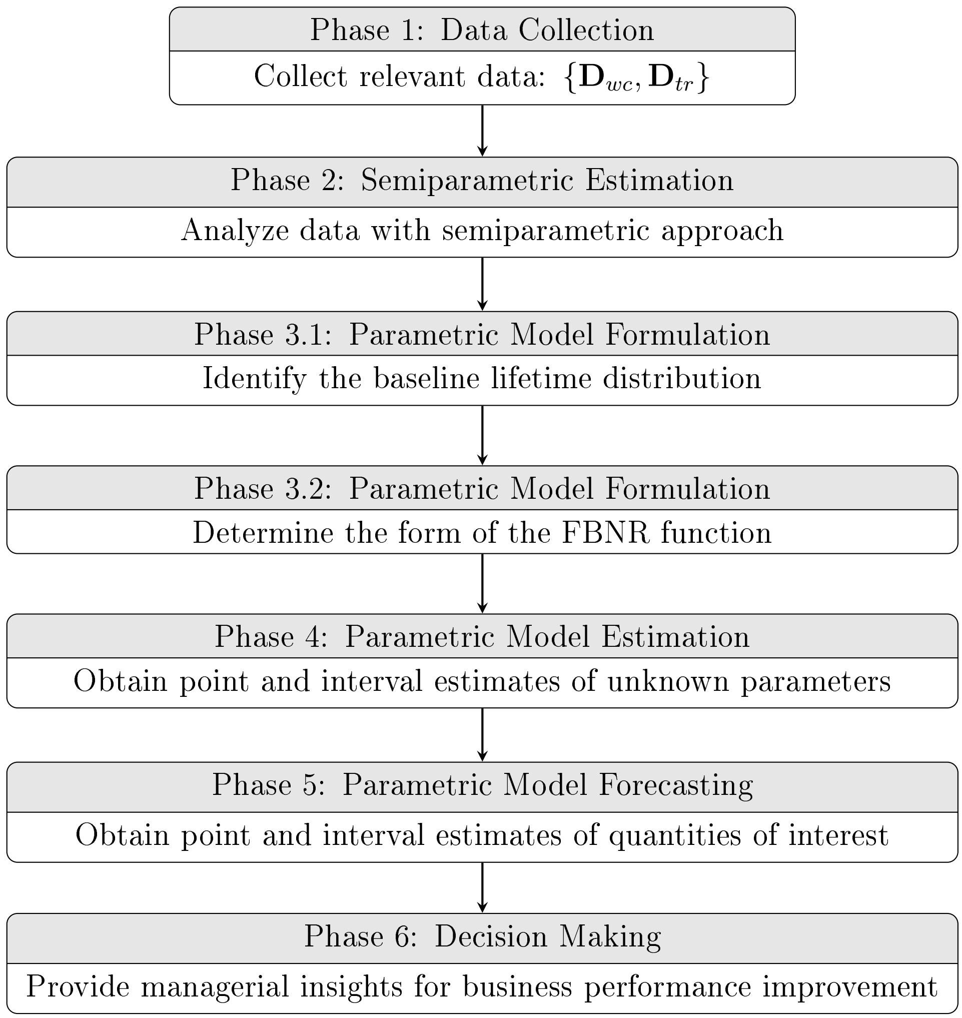


Figure 2: A framework for the statistical analysis and inference.

# 4 Semiparametric Estimation

Through joint analysis of tracking data and warranty claim data, this section develops semiparametric inference techniques by assuming a nonparametric lifetime CDF $F(\cdot)$ and a parametric FBNR function. We first consider a general FBNR function given in (1). A weighted estimator of $F(\cdot)$ is proposed. We then consider a special case when the FBNR proportion is a constant, where the closed-form of the weight that minimizes the variance of the weighted estimator of $F(\cdot)$ is available.

### 4.1 General FBNR function

Assume that $q(\cdot)$ takes the parametric form in (1). For $n$ untracked units, suppose that there are $H$ ($H \leq n$) distinct failure times $t_1 < t_2 < \cdots < t_H$. let $d_h = \sum_{i=1}^{n} I(t_i = t_h, \delta_i = 1)$ be the number of failures at time $t_h$, and $n_h = \sum_{i=1}^{n} I(t_i \geq t_h)$ be the number of units at risk just prior to $t_h$. Similarly, for $m$ tracked units, suppose that there are $K$ distinct failure times $t_1 < t_2 < \cdots < t_K$. let $d_k = \sum_{j=1}^{m} I(t_j = t_k, \tilde{\delta}_j = 1)$ be the number of failures at time $t_k$, and $m_k = \sum_{j=1}^{m} I(t_j \geq t_k)$ be the number of units at risk just prior to $t_k$. Let $G(\cdot)$ be the failure-report proportion of the untracked units, and $F(\cdot)$ is the product lifetime distribution. The Kaplan-Meier (KM) estimates of $G(\cdot)$ and $F(\cdot)$ can be readily obtained, respectively:

$$\widehat{G}(t) = 1 - \prod_{h:\, t_h \leq t} \left(1 - \frac{d_h}{n_h}\right), \tag{2}$$

and

$$\widehat{F}(t) = 1 - \prod_{k:\, t_k \leq t} \left(1 - \frac{d_k}{m_k}\right). \tag{3}$$

It is easy to know that $G(t) = \int_0^t f(x)[1 - q(x)]dx$. Apply integration by parts to get

$$G(t) = F(t)\left[1 - q(t)\right] + \int_0^t F(x) q'(x) dx. \tag{4}$$

Replacing $G(t)$ and $F(t)$ with their respective estimates, and noting that $\widehat{F}(t)$ is constant over $[t_{k-1}, t_k)$, we have the following estimating equations for the parameters in $q(\cdot)$:

$$\widehat{G}(t) = \widehat{F}(t_{k(t)})\left[1 - q(t_{k(t)})\right] + \sum_{k:t_k \leq t} \widehat{F}(t_{k-1})\left[q(t_k) - q(t_{k-1})\right], \tag{5}$$

where $k(t) = \max\{k : t_k \leq t\}$ and $t_0 = 0$. Evaluating the above equation at three separate points and solving the system of equations, we can obtain estimates of the FBNR parameters $(r_0, a, b)$. According to our simulation experience, we suggest evaluating at the three time points $t = \tau/3$, $\tau/2$ and $\tau$ for simplicity, or the three points $0.15505\tau$, $0.64495\tau$ and $\tau$ inspired from the Legendre-Gauss Quadrature method (Canuto et al., 1988, sec. 2.3).

The uniform strong consistency of the KM estimator under random censorship ensures that $\sup_{t:0\le t\le\tau}|\widehat{F}(t) - F(t)| \to 0$ as $m \to \infty$ and $\sup_{t:0\le t\le\tau}|\widehat{G}(t) - G(t)| \to 0$ as $n \to \infty$ (Stute and Wang, 1993). If $(r_0, a, b)$ is the unique solution of the system of equations:

$$G(\tau_i) = F(\tau_i)\left[1 - q(\tau_i)\right] + \int_0^{\tau_i} F(x)q'(x)dx, i = 1, 2, 3,$$

then we may treat $(\widehat{r}_0, \widehat{a}, \widehat{b})$ as $Z$-estimators. Therefore, $(\widehat{r}_0, \widehat{a}, \widehat{b}) \to_{a.s.} (r_0, a, b)$ as $n, m \to \infty$ by using Theorem 5.9 in Van der Vaart (2000) together with a continuous mapping argument applying to the integral of $F(x)q'(x)$. Based on the standard theory of KM estimation, both $\widehat{G}(t)$ and $\widehat{F}(t)$ are asymptotically normal with the respective asymptotic variance as

$$\sigma^2_{\widehat{G}}(t) = \mathrm{AVar}\{\sqrt{n}[\widehat{G}(t) - G(t)]\} = [1 - G(t)]^2 \int_0^t \frac{dG(x)}{[1 - G(x)]^2[1 - C(x)]}, \tag{6}$$

and

$$\sigma^2_{\widehat{F}}(t) = \mathrm{AVar}\{\sqrt{m}[\widehat{F}(t) - F(t)]\} = [1 - F(t)]^2 \int_0^t \frac{dF(x)}{[1 - F(x)]^2[1 - C(x)]}, \tag{7}$$

where $C(\cdot)$ is the censoring time distribution that is assumed common for both datasets. The estimates of $\sigma^2_{\widehat{G}}(t)$ and $\sigma^2_{\widehat{F}}(t)$ are given by

$$\widehat{\sigma}^2_{\widehat{G}}(t) = n\left[1 - \widehat{G}(t)\right]^2 \sum_{h:\, t_h \le t} \frac{d_h}{n_h(n_h - d_h)}, \tag{8}$$

and

$$\widehat{\sigma}^2_{\widehat{F}}(t) = m\left[1 - \widehat{F}(t)\right]^2 \sum_{k:\, t_k \le t} \frac{d_k}{m_k(m_k - d_k)}, \tag{9}$$

respectively (Lawless, 2003, pp. 97).

Although the KM estimator $\widehat{F}(t)$ from tracking data has good large sample properties, the size of the tracking data is usually not large as tracking could be costly. After obtaining $\widehat{G}(t)$ and the estimated $\widehat{q}(t) = 1 - \widehat{r}_0 \exp\left[-(\widehat{a}t)^{\widehat{b}}\right]$, an alternative estimator of $F(t)$ can be obtained recursively based on (4), denoted as $\widetilde{F}(t)$. Noting that $\widetilde{F}(t)$ is constant over $[t_{h-1}, t_h)$, we have the following equation:

$$\widehat{G}(t_{h(t)}) = \widetilde{F}(t_{h(t)}) \left[1 - \widehat{q}(t_{h(t)})\right] + \sum_{h:t_h \leq t} \widetilde{F}(t_{h-1}) \left[\widehat{q}(t_h) - \widehat{q}(t_{h-1})\right], \tag{10}$$

where $h(t) = \max\{h : t_h \leq t\}$ and $t_0 = 0$. Therefore, we have

$$\widetilde{F}(t) = \widetilde{F}(t_{h(t)}) = \frac{\widehat{G}(t_{h(t)}) - \sum_{h:t_h \leq t} \widetilde{F}(t_{h-1}) \left[\widehat{q}(t_h) - \widehat{q}(t_{h-1})\right]}{1 - \widehat{q}(t_{h(t)})}. \tag{11}$$

For example, when $t = t_1$, we have $\widetilde{F}(t_1) = \widehat{G}(t_1)/[1 - \widehat{q}(t_1)]$.

To take advantage of both tracking data and warranty claim data, we may pool $\widehat{F}(t)$ and $\widetilde{F}(t)$ to get a weighted estimator of $F(t)$ as

$$\widehat{F}^*(t; w) = w\widetilde{F}(t) + (1 - w)\widehat{F}(t), \tag{12}$$

where $w$ is a weight parameter. A rule of thumb choice for the weight is $w = n/(n+m)$. We may also consider a choice of $w$ that minimizes $\mathrm{AVar}(\widehat{F}^*(t; w))$. The minimization entails a closed form of $\mathrm{AVar}(\widehat{F}^*(t; w))$, which is not available when a general FBNR function (1) is assumed. When the FBNR proportion is constant, nevertheless, a closed form of such a $w$ is available, as shown in the next subsection.

## 4.2 Constant FBNR function

When $a$ is zero in (1), the FBNR function is a constant function, i.e., $q(t) = 1 - r_0$. As a result, we will have $G(t) = r_0 F(t)$. For any $t \leq \tau$, an estimator of $r_0$ and its corresponding asymptotic variance can be obtained as

$$\widehat{r}_0(t) = \frac{\widehat{G}(t)}{\widehat{F}(t)}, \tag{13}$$

and

$$\sigma^2_{\widehat{r}_0}(t) = \frac{\sigma^2_{\widehat{G}}(t)}{F(t)^2} + \frac{r_0^2 \sigma^2_{\widehat{F}}(t)}{F(t)^2}. \tag{14}$$

The choice of $t$ in $\widehat{r}_0(t)$ above affects the asymptotic variance. We recommend using $t = \tau$, i.e., the warranty length, for $\widehat{r}_0(t)$. To justify this choice, we show that $\sigma^2_{\widehat{r}_0}(t)$ is decreasing with $t$ in a special case where censoring times of all units are equal to $\tau$. See Appendix C of the online supplementary material for more details.

Based on $\widehat{r}_0^* = \widehat{G}(\tau)/\widehat{F}(\tau)$, the estimator $\widetilde{F}(t)$ from the warranty claim data and the weighted estimator $\widehat{F}^*(t;w)$ based on all data are then given by

$$\widetilde{F}(t) = \frac{\widehat{G}(t)\widehat{F}(\tau)}{\widehat{G}(\tau)}, \tag{15}$$

and

$$\widehat{F}^*(t;w) = w\frac{\widehat{G}(t)\widehat{F}(\tau)}{\widehat{G}(\tau)} + (1-w)\widehat{F}(t), \tag{16}$$

respectively. The asymptotic variance of $\widehat{F}^*(t;w)$ is given by

$$\sigma^2_{\widehat{F}^*}(t;w) = \frac{w^2\left[\eta_1 G(\tau)^2 - 2\eta_2 G(\tau)G(t) + \eta_3 G(t)^2\right]}{G(\tau)^4} - \frac{2\eta_4 w}{G(\tau)} + \eta_5, \tag{17}$$

and by solving the equation $\partial\sigma^2_{\widehat{F}^*}(t;w)/\partial w = 0$, we can obtain the optimum weight $w^*$ as

$$w^* = \frac{\eta_4 G(\tau)^3}{\eta_1 G(\tau)^2 - 2\eta_2 G(\tau)G(t) + \eta_3 G(t)^2},$$

where $\eta_1 = F(\tau)^2\sigma^2_{\widehat{G}}(t) + G(\tau)^2\sigma^2_{\widehat{F}}(t)$, $\eta_2 = F(\tau)^2\sigma_{\widehat{G}}(t,\tau) + G(\tau)^2\sigma_{\widehat{F}}(t,\tau)$, $\eta_3 = F(\tau)^2\sigma^2_{\widehat{G}}(\tau) + G(\tau)^2\sigma^2_{\widehat{F}}(\tau)$, $\eta_4 = G(\tau)\sigma^2_{\widehat{F}}(t) - G(t)\sigma_{\widehat{F}}(t,\tau)$, $\eta_5 = \sigma^2_{\widehat{F}}(t)$. Note that $\sigma_{\widehat{G}}(t,\tau)$ and $\sigma_{\widehat{F}}(t,\tau)$ are the respective covariance functions of the processes $\{\sqrt{n}[\widehat{G}(t) - G(t)], 0 < t \leq \tau\}$ and $\{\sqrt{m}[\widehat{F}(t) - F(t)], 0 < t \leq \tau\}$ that converge weekly to zero mean Gaussian processes (Lawless, 2003, pp. 97). For $t \leq \tau$, $\sigma_{\widehat{G}}(t,\tau)$ and $\sigma_{\widehat{F}}(t,\tau)$ can be obtained as

$$\sigma_{\widehat{G}}(t,\tau) = [1-G(t)][1-G(\tau)]\int_0^t \frac{dG(x)}{[1-G(x)]^2[1-C(x)]}, \tag{18}$$

and

$$\sigma_{\widehat{F}}(t,\tau) = [1-F(t)][1-F(\tau)]\int_0^t \frac{dF(x)}{[1-F(x)]^2[1-C(x)]}. \tag{19}$$

The respective estimates of $\sigma_{\widehat{G}}(t,\tau)$ and $\sigma_{\widehat{F}}(t,\tau)$ are given by

$$\widehat{\sigma}_{\widehat{G}}(t,\tau) = n\left[1-\widehat{G}(t)\right]\left[1-\widehat{G}(\tau)\right]\sum_{h:\, t_h \leq t} \frac{d_h}{n_h(n_h - d_h)}, \tag{20}$$

and

$$\widehat{\sigma}_{\widehat{F}}(t,\tau) = m\left[1-\widehat{F}(t)\right]\left[1-\widehat{F}(\tau)\right]\sum_{k:\, t_k \leq t} \frac{d_k}{m_k(m_k - d_k)}. \tag{21}$$

Based on these estimates, an estimate of the optimal weight $w^*$ can be obtained.

# 5 Parametric Inference

The above section has discussed semiparametric inference that assumes a parametric FBNR function (1) and a nonparametric lifetime distribution $F(\cdot)$. The nonparametric estimate of $F(\cdot)$ together with a probability plot is helpful to identify an appropriate distribution in an appropriate parametric distribution, based on which parametric inference can be conducted.

## 5.1 Parametric lifetime model

The log-location-scale family of distributions is usually a good model for product lifetime (Lawless, 2003, chap. 1). Assume that the random lifetime $X$ belongs to the log-location-scale distribution with a cumulative distribution function (CDF)

$$F(x) = \Phi\left[\frac{\log(x) - \mu}{\sigma}\right], \qquad x > 0, \tag{22}$$

and a probability density function (PDF)

$$f(x) = \frac{1}{\sigma x}\phi\left[\frac{\log(x) - \mu}{\sigma}\right], \qquad x > 0, \tag{23}$$

respectively. The distribution has a location parameter $\mu \in \mathcal{R}$ and a scale parameter $\sigma > 0$. $\Phi(\cdot)$ and $\phi(\cdot)$ are the respective standard CDF and PDF corresponding to $\mu = 0$ and $\sigma = 1$. When $\phi(\cdot)$ is replaced by the standard normal PDF $\phi_{\text{nor}}(\cdot)$, the distribution is lognormal for $X$. While $\phi_{\text{sev}}(z) = \exp[z - \exp(z)]$ corresponds to a Weibull distribution for $X$, and $\phi_{\text{logis}}(z) = \exp(z)/[\sigma z(1 + \exp(z))^2]$ leads to a log-logistic distribution.

## 5.2 ML estimation

By assuming a log-location-scale distribution for the lifetime and a time-varying execution function $q(t)$ in (1), the parameters to estimate are $\boldsymbol{\theta} = (\mu, \sigma, r_0, a, b)'$. Maximum Likelihood (ML) estimation is used to estimate the parameters. First, consider the warranty claim data $\mathbf{D}_{wc}$. Because the report behavior is independent of the failure pattern, the likelihood contribution from a reported failure $(t_i, \delta_i)$ with $\delta_i = 1$ is equal to $[1 - q(t_i)]f(t_i)$. On the other hand, a unit is not reported due to two possibilities: (i) its lifetime is larger than the censoring time, or (ii) it has failed within warranty but the failure is not reported. Therefore, the likelihood contribution for an unreported record $(t_i, \delta_i)$ with $\delta_i = 0$ is $1 - F(t_i) + \int_0^{t_i} q(x)f(x)dx$, which is equal to $1 - \int_0^{t_i}[1 - q(x)]f(x)dx$.

Based on the above analysis, the log-likelihood function based on the warranty claim data $\mathbf{D}_{wc}$ is given by

$$\ell_{wc}(\boldsymbol{\theta}; \mathbf{D}_{wc}) = \sum_{i=1}^{n} \left\{ \delta_i \log\{[1 - q(t_i)]f(t_i)\} + (1 - \delta_i) \log \left\{ 1 - \int_0^{t_i} [1 - q(x)]f(x)dx \right\} \right\}. \tag{24}$$

Since we assume no FBNR phenomenon in the tracking data, the tracking data are simply right-censored and the associated log-likelihood function can be easily specified as:

$$\ell_{tr}(\boldsymbol{\theta}; \mathbf{D}_{tr}) = \sum_{j=1}^{m} \left\{ \tilde{\delta}_j \log[f(t_j)] + (1 - \tilde{\delta}_j) \log[1 - F(t_j)] \right\}. \tag{25}$$

Combining both the warranty claim data and the tracking data, we can get the joint log-likelihood function of $\boldsymbol{\theta}$, which is the sum of $\ell_{wc}(\boldsymbol{\theta}; \mathbf{D}_{wc})$ and $\ell_{tr}(\boldsymbol{\theta}; \mathbf{D}_{tr})$:

$$\begin{aligned} \ell(\boldsymbol{\theta}; \mathbf{D}_{wc}, \mathbf{D}_{tr}) = & \sum_{i=1}^{n} \left\{ \delta_i \log\{[1 - q(t_i)]f(t_i)\} + (1 - \delta_i) \log \left\{ 1 - \int_0^{t_i} [1 - q(x)]f(x)dx \right\} \right\} \\ & + \sum_{j=1}^{m} \left\{ \tilde{\delta}_j \log[f(t_j)] + (1 - \tilde{\delta}_j) \log[1 - F(t_j)] \right\}. \end{aligned} \tag{26}$$

The ML estimator $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\sigma}, \widehat{r}_0, \widehat{a}, \widehat{b})'$ can be obtained by numerically maximizing the log-likelihood function (26).

## 5.3 Confidence intervals

In addition to the point estimate, a confidence interval is commonly used to account for the uncertainty in the ML estimator. Both asymptotic normal approximation and bootstrap method will be used for the interval estimation.

Asymptotic normal approximation to the ML estimators $\widehat{\boldsymbol{\theta}}$ is often used to construct confidence intervals for $\boldsymbol{\theta}$. To use the method, we first obtain the ML estimates of $\widehat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$ through maximizing the joint log-likelihood function (26). Given a large sample of lifetime observations, the ML estimator $\widehat{\boldsymbol{\theta}}$ has a distribution which can be approximated by a multivariate normal distribution $N(\boldsymbol{\theta}, \mathcal{I}^{-1}(\boldsymbol{\theta}))$, where $\mathcal{I}(\boldsymbol{\theta})$ is the Fisher information matrix that can be estimated using the observed information matrix $I(\widehat{\boldsymbol{\theta}})$, the negative of the second-order partial derivatives of the joint log-likelihood function $\ell(\boldsymbol{\theta})$ evaluated at $\widehat{\boldsymbol{\theta}}$. Expressions of $I(\boldsymbol{\theta})$ when the FBNR function is a constant are provided in Appendix A of the online supplementary material. The method often works well for moderate sample sizes. The intervals so obtained are called normal approximation confidence intervals, which are easy to compute and widely used in many software packages (Jeng and Meeker, 2000). For a bounded parameter that takes value on a subset of the real line, normal approximation might generate intervals with endpoints outside the parameter space. Such absurd phenomenon could happen when the number of observed failures is not large enough or when the parameter value is close to the boundary. For example, the normal approximation confidence interval of $r_0 \in [0, 1]$ may have endpoints lying outside the parameter space when $r_0$ is close to zero or one. Consequently, a suitable reparameterization that transforms the parameter to a new scale may be needed to remove the range constraint on a bounded parameter. The logarithm transformation is chosen for non-negative parameters, e.g., the scale parameter $\sigma$. Two popular choices for $r_0 \in [0, 1]$ are the "probit" transformation $\Phi^{-1}_{\text{nor}}(r_0)$ and the "logit" transformation $\Phi^{-1}_{\text{logis}}(r_0) = \log[r_0/(1 - r_0)]$. According to our simulation studies, performances of these two transformations are almost the same.

Field data often have small number of observed failures and heavy censoring. In this case, the coverage probabilities of confidence intervals produced by the normal approximation method may not be close to nominal values (Jeng and Meeker, 2000). A carefully designed bootstrap procedure may be able to produce better confidence intervals. For our problem with random and heavy censoring, the parametric bootstrap may not be applicable because of the difficulties in specifying the censoring process. See Meeker and Escobar (1998, chap. 9) for more details. The nonparametric bootstrap can also be problematic as there is a high chance to get a bootstrap sample with no failures (Hong et al., 2009). In this paper, Newton and Raftery's weighted likelihood bootstrap is adopted (Newton and Raftery, 1994). The method uses a random weight to the log-likelihood of each observation and thus it is free of the resampling problems above. Given the warranty claim data $\mathbf{D}_{wc}$ and the tracking data $\mathbf{D}_{tr}$, the algorithm is as follows.

1. Simulate two sets of i.i.d. weight samples $\{w_i, i = 1, \cdots, n\}$ and $\{w_j, j = 1, \cdots, m\}$ from a probability distribution $F_W(\cdot)$.

2. Use the weights and two data sets to obtain a weighted log-likelihood function as

$$\widetilde{\ell}(\boldsymbol{\theta}; \mathbf{D}_{wc}, \mathbf{D}_{tr}) = \widetilde{\ell}_{wc}(\boldsymbol{\theta}; \mathbf{D}_{wc}) + \widetilde{\ell}_{tr}(\boldsymbol{\theta}; \mathbf{D}_{tr}), \tag{27}$$

   where

$$\widetilde{\ell}_{wc}(\boldsymbol{\theta}; \mathbf{D}_{wc}) = \sum_{i=1}^{n} w_i \left\{ \delta_i \log\{[1 - q(t_i)]f(t_i)\} + (1 - \delta_i) \log \left\{ 1 - \int_0^{t_i} [1 - q(x)]f(x)dx \right\} \right\},$$

   and

$$\widetilde{\ell}_{tr}(\boldsymbol{\theta}; \mathbf{D}_{tr}) = \sum_{j=1}^{m} w_j \left\{ \tilde{\delta}_j \log[f(t_j)] + (1 - \tilde{\delta}_j) \log[1 - F(t_j)] \right\}.$$

   Maximize the weighted log-likelihood to get an estimate $\widehat{\boldsymbol{\theta}}^*$.

3. Repeat the above steps for $B$ times to obtain a collection of bootstrap estimates $\widehat{\boldsymbol{\theta}}^{*(b)}, b = 1, 2, \cdots, B$.

In the weighted bootstrap, the randomness of $\widetilde{\ell}$ is induced by the random weights. For a weight distribution $F_W(\cdot)$ having the property $\mathbb{E}(W) = [\text{Var}(W)]^{1/2}$, Jin et al. (2001) showed that the conditional distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}}^* - \widehat{\boldsymbol{\theta}})$ given the data can provide a good approximation to the distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$. They also showed that the method is robust for different weight distributions, e.g, Gamma(1,1), Gamma(1,0.5), Beta($\sqrt{2} - 1, 1$). Based on the bootstrap estimates $\widehat{\theta}_i^{*(b)}, b = 1, 2, \cdots, B$ of the $i$th element of $\boldsymbol{\theta}$, the weighted bootstrap percentile confidence interval at a $100(1 - \alpha)\%$ confidence level is given by

$$[\widehat{\theta}^*_{i,\alpha/2}, \widehat{\theta}^*_{i,1-\alpha/2}],$$

where $\widehat{\theta}^*_{i,\alpha/2}$ and $\widehat{\theta}^*_{i,1-\alpha/2}$ are the $100(\alpha/2)$th and the $100(1-\alpha/2)$th quantile of the bootstrap estimates.

# 6 Goodness-of-Fit Test and Hypothesis Test

Before parameter estimation, goodness-of-fit tests and hypothesis tests in this section help identify appropriate parametric models for product's lifetime and customer's FBNR behavior. Statistically, models that provide best fits to the observed data will be chosen.

## 6.1 Select a baseline lifetime distribution

Prior knowledge is helpful in identifying a lifetime model. An appropriate lifetime model could be determined based on engineering knowledge and experience of experts. If the product is an upgraded version of an existing model, such as laptops and washing machines, then the lifetime distribution of the product may be the same as the old version. Alternatively, knowledge about the lifetime distribution may accumulate during the product development and lab testing. In these cases, the distribution assumption can be validated based on field data. On the other hand, it is common that the distribution is unclear to the manufacturer, due to a lack of knowledge. The field data can then help to select a baseline lifetime distribution. Qualitatively, the model selection can be done using a probability plot. A probability plot for a log-location-scale distribution compares the transformed nonparametric estimates of $F(\cdot)$ (using the transformation $\Phi^{-1}(\cdot)$ in (22)) with the logarithm of time. The KM estimates $\widehat{F}(\cdot)$ in (3) from right-censored tracking data might be used for the plot. While we suggest using the weighted estimates $\widehat{F}^{*}(\cdot)$ in (12), which are pooled estimates based on the tracking data and the warranty claim data, for the probability plot. Quantitatively, the Akaike information criterion (AIC), a statistic trading off a model's likelihood against its complexity, can be used for lifetime model selection. The AIC is defined as $\text{AIC} = -2\ell(\widehat{\boldsymbol{\theta}}) + 2\text{dim}(\boldsymbol{\theta})$, where $\text{dim}(\boldsymbol{\theta})$ is the number of model parameters. A model with a minimum AIC will be selected. The above qualitative and quantitative methods are demonstrated in Section 8.

Table 1: Discussion of hypothesis results.

| | | | | |
|---|---|---|---|---|
| $H_0 : a = 0$ | Reject | Reject | Not reject | Not reject |
| $H_0 : r_0 = 1$ | Reject | Not reject | Reject | Not reject |
| $q(t)$= | $1 - r_0 \exp\left[-(at)^b\right]$ | $1 - \exp\left[-(at)^b\right]$ | $1 - r_0$ | 0 |

### 6.2 Choose an appropriate FBNR function

After a parametric lifetime distribution is chosen, hypothesis testing can be used to determine a proper form of the FBNR function $q(t)$ based on the warranty claim data and the tracking data. A flexible FBNR model $q(t) = 1 - r_0 \exp\left[-(at)^b\right]$ has been proposed in Section 2.2. It is possible that some parameters take values at the boundary. To identify a suitable form for the function, we can test whether $a = 0$, or whether $r_0 = 1$. The two hypothesis tests can be constructed as follows:

$$H_0 : a = 0 \;\text{ versus }\; H_1 : a > 0,$$

and

$$H_0 : r_0 = 1 \;\text{ versus }\; H_1 : r_0 < 1.$$

The likelihood ratio test is used since the null model under $H_0$ could be viewed as a nested model of the alternative model under $H_1$. It should be noted that the parameter of interest is on the boundary of the parameter space, i.e., $a = 0$ or $r_0 = 1$. In this case, the asymptotic distribution of the likelihood ratio statistic is a $50 : 50$ mixture of $\chi_0^2$ and $\chi_1^2$ (Self and Liang, 1987). This fact can be used to test the hypothesis.

As shown in the Table 1, if both $H_0 : a = 0$ and $H_0 : r_0 = 1$ are rejected, the FBNR function is of the form $q(t) = 1 - r_0 \exp\left[-(at)^b\right]$, indicating a time-dependent increasing FBNR function. The parameter estimation is based on the log-likelihood function (26). If $H_0 : a_0 = 0$ is rejected but $H_0 : r_0 = 1$ is not, it indicates that there is no FBNR phenomenon at the beginning and an increasing FBNR proportion thereafter, i.e., $q(t) = 1 - \exp\left[-(at)^b\right]$. If $H_0 : r_0 = 1$ is reject but there is no sufficient evidence to reject $H_0 : a = 0$, we will conclude that the FBNR proportion is constant over time, i.e., $q(t) = 1 - r_0$.

If both $H_0 : a = 0$ and $H_0 : r_0 = 1$ are not rejected, we may accept the fact that all failures of untracked units will be reported, i.e., there is no FBNR phenomenon. Then both the tracking data and the warranty claim data are simply right censored.

# 7 Simulation Study

A simulation experiment is conducted to evaluate the performance of the proposed methods. In the simulation, the lifetime distribution is assumed lognormal with $\mu = 8.5$ and $\sigma = 1.5$, while the FBNR function is $q(t) = 1 - \exp\left[-(at)^b\right]$ with $a = 0.0013$ and $b = 10$. Both the warranty length $\tau$ and the end-of-study date are set to be 730. As 1000 tracked units are enough to achieve satisfactory performance according to our simulation experience, the tracking sample size is simply set to be 1000. Several untracked sample sizes are used, i.e., $n = 5000, 50000$.

Six distributions of the entry time, which is the time between shipment to a dealer and sale to a customer, were used to cover a variety of possible observational situations. Specifically, we generated entry times for both untracked sample and tracked sample from each of the uniform distribution on $(0, 30]$, $(0, 365]$, $(0, 730]$ and the exponential distribution with rate 1/15, 2/365, 1/365. The uniform distribution on $(0, 30]$ determines a staggered entry distribution with a mean of 15 days and a standard deviation of 8.7 days. The exponential distribution with rate 1/15 reflects a similar situation that most units are sold within one month. Other distributions of the entry time can be interpreted in similar ways. For each combination of $n$ and entry types, we generate 5000 Monte Carlo samples. Table 2 shows the estimated relative biases (bias divided by true value) and relative RMSEs (RMSE divided by true value) of parameter estimators given different sizes of warranty claim data and types of entry. When there are 5000 untracked units in total, both relative biases and RMSEs increase when the mean time of entry increases. This is because the missing data rate increase greatly as the observation window for a unit sold late is much shorter than a unit sold early. Another similar observation is that given the same mean time to entry, the relative biases and RMSEs are smaller with exponentially distributed entry times compared with that with entry times from a uniform distribution. A possible explanation is that the exponential distribution is skewed right so that more units will enter early with longer observation periods. For all entry types, nevertheless, the biases and RMSEs can be reduced significantly if the untracked sample size increases to 50000. As can be seen, with 50000 untracked units, the ML estimators have negligible biases and the RMSEs are reasonably small.

Besides staggered entry patterns, we also examined effects of warranty length and FBNR proportion. Without loss of generality, we considered the situation that all units are sold at time zero. In this case, the data are singly censored at the end of warranty. Table 3 shows the relative biases and relative RMSEs of parameter estimates with different warranty lengths. Note that the end-of-study date is equal to the warranty length for each case. As can be seen, the relative biases are reasonably small and the relative RMSEs decrease with longer warranty since missing data rate decreases. Different FBNR proportions are achieved by varying $b$. As we fixed $a$ as 0.0013 and warranty length as well as end-of-study date as 730. Higher value of $b$ results in lower FBNR proportion. Table 4 shows the estimated relative biases and relative RMSEs of parameter estimators with different values of $b$. As can be seen, the RMSEs decrease with the increase of $b$ as the overall missing data rate decreases.

As shown in the simulation, the inference accuracy of estimators of FBNR parameters is greatly affected by the size of the warranty data. The missing data proportion, which is essentially determined by staggered entry types, length of warranty, end-of-study date, and the FBNR function, also has great impacts on the estimation accuracy and efficiency. In reality, the warranty sample is usually very large. Together with relatively small number of tracking units, we are able to decouple FBNR information by joint analyzing the warranty data and the tracking data. With more warranty claim data, estimates will be more accurate and useful. The simulation results provide useful insights on warranty data collection.

Table 2: Estimated relative biases and relative RMSEs of parameter estimators with different warranty claim data sizes and different staggered entry types.

| $n$ | Parameter | Unif(0,30) | | Unif(0,365) | | Unif(0,730) | |
|---|---|---|---|---|---|---|---|
| | | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) |
| 5000 | $\mu$ | -0.03 | 1.11 | -0.05 | 1.29 | -0.15 | 1.61 |
| | $\sigma$ | -0.14 | 3.89 | -0.21 | 4.27 | -0.41 | 5.00 |
| | a | -1.37 | 10.40 | -7.75 | 27.51 | -10.33 | 32.67 |
| | b | 5.41 | 37.97 | 5.50 | 63.78 | 6.01 | 85.35 |
| 50000 | $\mu$ | 0.02 | 0.39 | 0.01 | 0.44 | 0.00 | 0.56 |
| | $\sigma$ | 0.02 | 1.35 | 0.01 | 1.46 | 0.01 | 1.74 |
| | a | 0.32 | 1.54 | -0.15 | 6.24 | -1.97 | 12.92 |
| | b | 3.19 | 5.65 | 4.03 | 6.23 | 4.33 | 6.82 |
| $n$ | Parameter | Exp(1/15) | | Exp(2/365) | | Exp(1/365) | |
| | | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) |
| 5000 | $\mu$ | -0.02 | 1.11 | -0.07 | 1.27 | -0.08 | 1.47 |
| | $\sigma$ | -0.06 | 3.89 | -0.30 | 4.24 | -0.22 | 4.76 |
| | a | -1.15 | 10.08 | -6.21 | 23.38 | -8.67 | 28.80 |
| | b | 4.61 | 22.73 | 4.55 | 30.03 | 5.34 | 54.22 |
| 50000 | $\mu$ | 0.03 | 0.39 | 0.02 | 0.44 | 0.01 | 0.50 |
| | $\sigma$ | 0.06 | 1.35 | 0.06 | 1.48 | 0.03 | 1.64 |
| | a | 0.26 | 1.51 | 0.16 | 3.20 | -0.37 | 6.74 |
| | b | 3.17 | 5.61 | 3.96 | 6.40 | 4.12 | 6.37 |

Table 3: Estimated relative biases and relative RMSEs of parameter estimators with different warranty lengths.

| $n$ | Parameter | $\tau = 730$ | | $\tau = 800$ | | $\tau = 900$ | |
|---|---|---|---|---|---|---|---|
| | | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) |
| 5000 | $\mu$ | 0.00 | 1.11 | -0.03 | 1.05 | 0.03 | 0.97 |
| | $\sigma$ | -0.02 | 3.90 | -0.11 | 3.76 | 0.03 | 3.53 |
| | a | -0.59 | 7.42 | 0.13 | 2.16 | 0.05 | 1.43 |
| | b | 4.29 | 18.46 | 3.97 | 15.83 | 3.28 | 11.76 |
| 50000 | $\mu$ | 0.02 | 0.39 | 0.05 | 0.34 | 0.04 | 0.33 |
| | $\sigma$ | 0.06 | 1.34 | 0.15 | 1.24 | 0.13 | 1.20 |
| | a | 0.24 | 1.31 | 0.04 | 0.67 | -0.07 | 0.42 |
| | b | 3.26 | 5.91 | 2.91 | 5.72 | 1.58 | 3.99 |

Table 4: Estimated relative biases and relative RMSEs of parameter estimators with different values of $b$.

| $n$ | Parameter | $b = 8$ | | $b = 10$ | | $b = 12$ | |
|---|---|---|---|---|---|---|---|
| | | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) | Biases(%) | RMSEs(%) |
| 5000 | $\mu$ | -0.03 | 1.17 | 0.00 | 1.11 | -0.03 | 1.05 |
| | $\sigma$ | -0.09 | 4.02 | -0.02 | 3.90 | -0.13 | 3.75 |
| | a | -0.32 | 6.11 | -0.59 | 7.42 | -0.82 | 8.28 |
| | b | 5.30 | 35.40 | 4.29 | 18.46 | 4.39 | 11.50 |
| 50000 | $\mu$ | 0.03 | 0.41 | 0.02 | 0.39 | 0.02 | 0.37 |
| | $\sigma$ | 0.08 | 1.41 | 0.06 | 1.34 | 0.04 | 1.30 |
| | a | 0.19 | 1.51 | 0.24 | 1.31 | 0.23 | 1.21 |
| | b | 2.80 | 7.49 | 3.26 | 5.91 | 3.41 | 5.78 |

# 8 Illustrative Example

Our example concerns an appliance that is used in residences and offices. The appliance, referred to as Product D in an early study (Hong and Meeker, 2010), had multiple failure modes and a two-year warranty. Some product units are installed with sensors and connected to the Internet. Their failure information is well tracked. For untracked units, the failure time information is available for those failed and returned to the manufacturer. They may suffer from the FBNR problem. The data were multiply censored due to the staggered entry of a product into the field over one month. To protect sensitive and proprietary information of the product, we simulate data based on the results from analysis of the real data. Without loss of generality, we focus on one specific failure mode of the Product D.

There are 1105 tracked units and 5126 untracked units in total. In the tracked group, 115 failures are observed. In the untracked group, 490 failures are reported and the other 4636 units do not claim a warranty. The respective KM estimates $\widehat{G}(t)$ and $\widehat{F}(t)$ of the report fractions for the untracked and the tracked units as well as the weighted estimates $\widehat{F}^{*}(t)$ based on all data are shown in Figure 3. As we can see, $\widehat{G}(t)$ for the untracked units are lower compared to $\widehat{F}(t)$ for the tracked units as well as $\widehat{F}^{*}(t)$, and the difference between the two proportions becomes larger over time. The observation indicates the existence of the FBNR behavior.

The lognormal, Weibull and loglogistic models were used to fit the tracking data and the resulting AIC values are 2233.5, 2242.1 and 2240.9 respectively. The lognormal distribution is selected for the data as it has the smallest AIC value. Figure 4 shows the lognormal probability plot based on $\widehat{F}^{*}(t)$. The linear pattern in the plot shows that the lognormal model provides a good fit to the data. To select an appropriate FBNR function, we proceed to test the hypothesis $H_0 : a = 0$ first. The likelihood ratio test shows that $H_0$ is rejected at a significance level of 0.05. While there is no sufficient evidence to reject the hypothesis $H_0 : r_0 = 1$ at a significance level of 0.05. Therefore, we use $q(t) = 1 - \exp\left[-(at)^b\right]$ for the FBNR function.

After the lifetime model and the FBNR function are determined, the parameter esti-

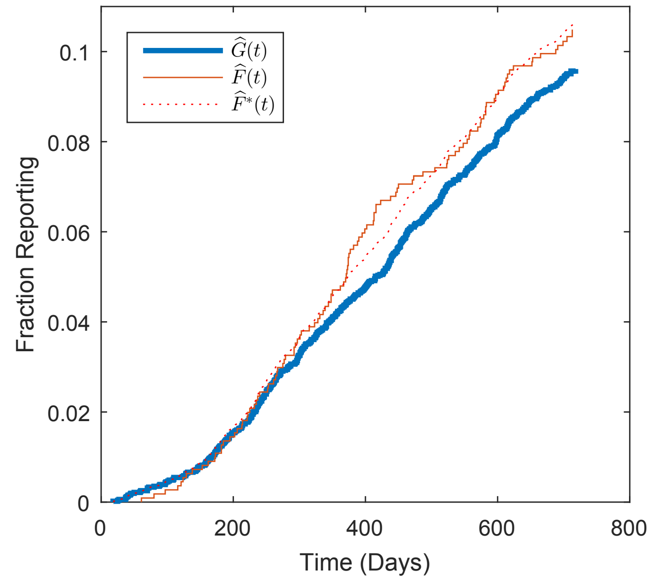


Figure 3: The respective KM estimates $\widehat{G}(t)$ and $\widehat{F}(t)$ of the report fractions for the untracked and the tracked units as well as the weighted estimates $\widehat{F}^*(t)$ based on all data.

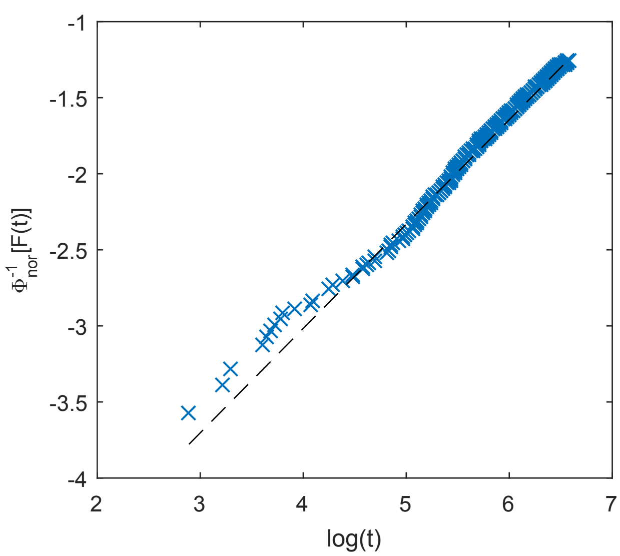


Figure 4: Lognormal probability plot based on $\widehat{F}^*(t)$.

Table 5: The ML estimates, standard errors and 95% approximate confidence intervals constructed by normal approximation (NOR) and weighted bootstrap percentile (WBP) based on warranty claim and tracking data together.

| Parameter | Estimate | Standard error | 95% Confidence interval | |
|---|---|---|---|---|
| | | | NOR | WBP |
| $\mu$ | 8.44 | 0.11 | [8.23, 8.65] | [8.15, 8.60] |
| $\sigma$ | 1.47 | 0.06 | [1.35, 1.60] | [1.32, 1.57] |
| $a$ | 0.0013 | 0.00005 | [0.0012, 0.0014] | [0.0007, 0.0014] |
| $b$ | 9.69 | 3.96 | [1.93, 17.45] | [1.41, 14.94] |

mates are obtained based on the tracking data and warranty claim data together. Table 5 gives the ML estimates, standard errors and the 95% approximate confidence intervals for the model parameters constructed by the normal approximation and the weighted bootstrap percentile method. As shown in the table, the standard errors of most estimates are relatively small except for that of $\widehat{b}$. This estimation problem is mainly caused by the heavy censoring and limited warranty claim data. Note that the multiple censoring causes further impreciseness of the inference when compared with that of single right censoring.

Figure 5 shows the ML estimates and the corresponding 95% pointwise confidence bands for $F(t)$ based on all data $\mathbf{D}_{wc} \cup \mathbf{D}_{tr}$ and based on tracking data $\mathbf{D}_{tr}$ only. The weighted bootstrap is used to obtain the confidence bands. As we can see, the confidence bands based on all data are narrower than that based on tracking data. Figure 6 shows the ML estimates and the corresponding 90% pointwise confidence bands for $q(t)$ using normal approximation with transformation based on all data. The FBNR proportion is close to zero and relatively low at the beginning and increases afterwards, which results in lower return proportion of untracked units than that of tracked units. In reality, manufacturers are suggested to pay special attention to the existence of the FBNR phenomenon and keep updating analyses with the accumulation of warranty claim data.

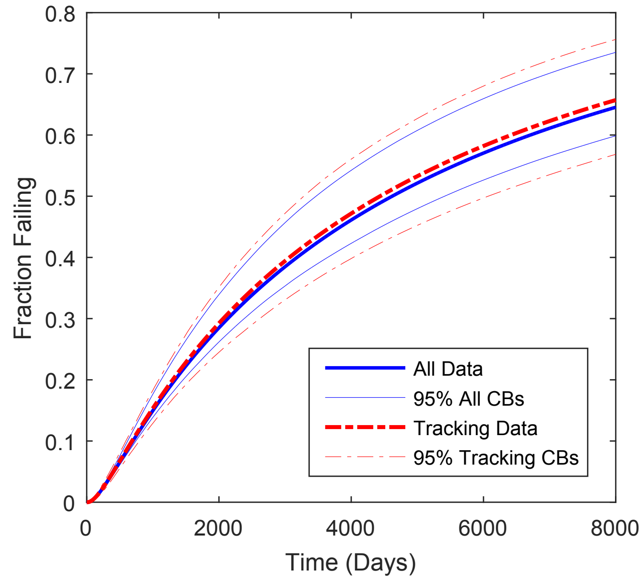


Figure 5: The ML estimates and corresponding 95% pointwise confidence bands (CBs) for $F(t)$ using weighted bootstrap percentile based on all data $\mathbf{D}_{wc} \cup \mathbf{D}_{tr}$ (solid line) and only based on tracking data $\mathbf{D}_{tr}$ (dash-dot line).

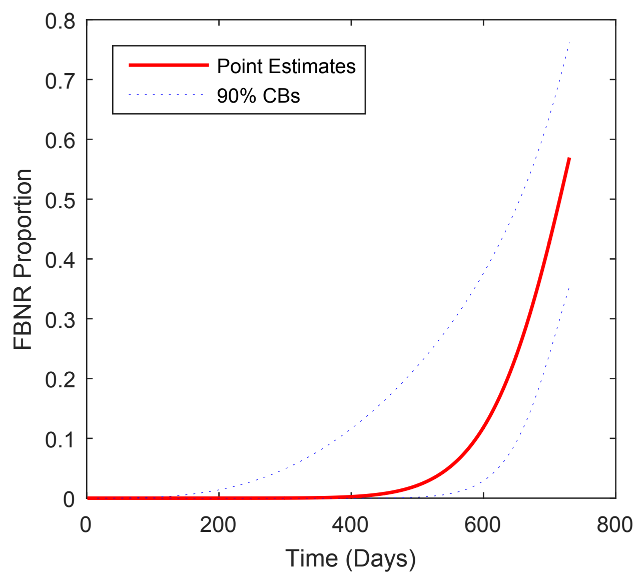


Figure 6: The ML estimates and corresponding 90% pointwise confidence bands (CBs) for $q(t)$ based on all data $\mathbf{D}_{wc} \cup \mathbf{D}_{tr}$.

# 9 Conclusion

The FBNR behavior is common due to several factors, such as the product price, the product life cycle, the warranty policy, and the quality of the warranty service. Ignorance of the FBNR phenomenon leads to an overestimate of product reliability based on field return data or an overestimate of warranty cost based on lab data or tracking data. In this study, we used both warranty claim data and tracking data for statistical inference in the presence of the FBNR problem. We proposed a flexible FBNR function (1) to model the time-varying FBNR behavior. Based on the FBNR function, a statistical framework was proposed to jointly analyze the warranty claim data and the tracking data. In the framework, we first performed semiparametric data analysis. Such analysis provides insights into the FBNR behavior and helps to identify an appropriate lifetime distribution for parametric analysis. After selecting an appropriate lifetime model and a FBNR function through goodness-of-fit and hypothesis testing, both point and interval estimations of the model parameters can be performed based on standard likelihood methods. The estimated FBNR proportion can then be used to predict warranty costs and spare-part provision of the current product. It can also be used for the prediction of a new version of the current product or a similar product. More importantly, the FBNR proportion can be used as an important indicator of customer satisfaction, which helps manufactures understand customer sentiments.

## SUPPLEMENTARY MATERIALS

The supplementary materials of this article include the following: (1) details and illustrations of the likelihood inference; (2) proof of the decreasing function (14) in Section 4.2; (3) R code for evaluating the performance of the proposed method.

## ACKNOWLEDGMENTS

We are grateful to the editor, associate editor, and two referees for their insightful comments and constructive suggestions. Ye Zhi-Sheng is supported by Singapore AcRF Tier 1 funding (#R-266-000-081-133 and #R-266-000-095-112), and also the Natural Science Foundation of China (71601138). The work by Hong Yili was partially supported by the National Science Foundation under Grant CMMI-1634867 to Virginia Tech.

# References


Blischke, W. R., Karim, M. R., and Murthy, D. N. P. (2011). *Warranty Data Collection and Analysis*. Springer.

Canuto, C., Hussaini, M. Y., Quarteroni, A., and Zang, T. A. (1988). *Spectral Methods in Fluid Dynamics*. Springer.

Chukova, S. and Johnston, M. (2006). Two-dimensional warranty repair strategy based on minimal and complete repairs. *Mathematical and Computer Modelling*, 44(11-12):1133–1143.

Gurgur, C. Z. (2011). Dynamic cash management of warranty reserves. *The Engineering Economist*, 56(1):1–27.

Hong, Y. and Meeker, W. Q. (2010). Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics*, 52(2):148–159.

Hong, Y., Meeker, W. Q., and McCalley, J. D. (2009). Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics*, 3(2):857–879.

Hsu, N.-J., Tseng, S.-T., and Chen, M.-W. (2015). Adaptive warranty prediction for highly reliable products. *IEEE Transactions on Reliability*, 64(3):1057–1067.

Hu, X. J. and Lawless, J. F. (1996). Estimation from truncated lifetime data with supplementary information on covariates and censoring times. *Biometrika*, 83(4):747–761.

Huang, H. Z., Liu, Z. J., and Murthy, D. N. P. (2007). Optimal reliability, warranty and price for new products. *IIE Transactions*, 39(8):819–827.

Ja, S.-S., Kulkarni, V. G., Mitra, A., and Patankar, J. G. (2002). Warranty reserves for nonstationary sales processes. *Naval Research Logistics*, 49(5):499–513.

Jeng, S. L. and Meeker, W. Q. (2000). Comparisons of approximate confidence interval procedures for type I censored data. *Technometrics*, 42(2):135–148.

Jin, Z., Ying, Z., and Wei, L. (2001). A simple resampling method by perturbing the minimand. *Biometrika*, 88(2):381–390.

Kalbfleisch, J. D. and Lawless, J. F. (1992). Some useful statistical methods for truncated data. *Journal of Quality Technology*, 24(3):145–152.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*. John Wiley & Sons.

Meeker, W. Q. and Escobar, L. A. (1998). *Statistical Methods for Reliability Data*. John Wiley & Sons.

Nair, V. N., Tang, B., and Xu, L.-A. (2001). Bayesian inference for some mixture problems in quality and reliability. *Journal of Quality Technology*, 33(1):16–28.

Newton, M. A. and Raftery, A. E. (1994). Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B (Methodological)*, 56(1):3–48.

Patankar, J. G. and Mitra, A. (1995). Effects of warranty execution on warranty reserve costs. *Management Science*, 41(3):395–400.

Rai, B. and Singh, N. (2006). Customer-rush near warranty expiration limit, and nonparametric hazard rate estimation from known mileage accumulation rates. *IEEE Transactions on Reliability*, 55(3):480–489.

Self, S. G. and Liang, K.-Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association*, 82(398):605–610.

Stevens, R. E. (2006). *The Marketing Research Guide*. Routledge.

Stute, W. and Wang, J.-L. (1993). The strong law under random censorship. *The Annals of Statistics*, 21(3):1591–1607.

Taylor, L. L. and Peña, E. A. (2014). Nonparametric estimation with recurrent competing risks data. *Lifetime Data Analysis*, 20(4):514–537.

Van der Vaart, A. W. (2000). *Asymptotic Statistics*. Cambridge University Press.

Wilhelm, W., Yankov, A., and Magee, P. (2011). Mobile phone consumption behavior and the need for sustainability innovations. *Journal of Strategic Innovation and Sustainability*, 7(2):20–40.

Wu, S. (2011). Warranty claim analysis considering human factors. *Reliability Engineering & System Safety*, 96(1):131–138.

Wu, S. (2013). A review on coarse warranty data and analysis. *Reliability Engineering & System Safety*, 114:1–11.

Xie, W. and Liao, H. (2013). Some aspects in estimating warranty and post-warranty repair demands. *Naval Research Logistics*, 60(6):499–511.

Xie, W., Liao, H., and Zhu, X. (2014). Estimation of gross profit for a new durable product considering warranty and post-warranty repairs. *IIE Transactions*, 46(2):87–105.

Ye, Z.-S., Hong, Y., and Xie, Y. (2013). How do heterogeneities in operating environments affect field failure predictions and test planning? *The Annals of Applied Statistics*, 7(4):2249–2271.