Product Component Genealogy Modeling and Field-Failure Prediction

Caleb King, Yili Hong, and William Q. Meeker

Abstract

Many industrial products consist of multiple components that are necessary for system operation. There is an abundance of literature on modeling the lifetime of such components through competing risks models. During the life-cycle of a product, it is common for there to be incremental design changes to improve reliability, to reduce costs, or due to changes in availability of certain part numbers. These changes can affect product reliability, but are often ignored in system lifetime modeling. By incorporating this information about changes in part numbers over time (information that is readily available in most product production databases), better accuracy can be achieved in predicting time to failure, thus yielding more accurate field-failure predictions. This paper presents methods for estimating parameters and predictions for this generational model and a comparison with existing methods through the use of simulation. Our results indicate that the generational model has important practical advantages and outperforms the existing methods in predicting field failures.

Index Terms

Competing risks, Component generation, Lifetime data, Multiple failure modes, Prediction intervals, Product reliability.

C. King and Y. Hong are with the Department of Statistics, Virginia Tech, Blacksburg, VA, 24061, USA (e-mail: calebk6@vt.edu; yilihong@vt.edu)

William Q. Meeker is with the Department of Statistics, Iowa State University, Ames, IA, 50011, USA (email: wqmeeker@iastate.edu)

ACRONYMS

cdf	cumulative distribution function
CI	confidence interval
DFD	data freeze date
ELCG	extended location-change generational
LCG	location-change generational
ML	maximum likelihood
MSE	mean square error
pdf	probability density function
PI	prediction interval
RMSE	relative mean square error
SCR	standard competing risks

NOTATION

n	total sample size of system units
T	failure time random variable
t_i	observed failure or censoring time for system i
J	total number of components in product
τ_i	installation time on the calendar time scale of system i
g_{ij}	generation information function of component j of system i
G_j	maximum number of generations for component j
δ_{ij}	observed failure indicator for component j of system i
Δ_{ij}	failure indicator for component j of system i
F_{ij}	cdf of failure time for system i failing due to component j
f_{ij}	pdf of failure time for system i failing due to component j
F^*_{ij}	sub-distribution function of system i failing due to component j
f^*_{ij}	sub-density function of system i failing due to component j
Φ_{sev}	standard smallest extreme value cdf
ϕ_{sev}	standard smallest extreme value pdf
Φ_{nor}	standard normal cdf
ϕ_{nor}	standard normal pdf
μ	location parameter of the distribution of $\log(T)$
σ	scale parameter of the distribution of $\log(T)$
$L(\boldsymbol{\theta}\|\text{DATA})$	likelihood function for the unknown parameter vector $\boldsymbol{\theta}$
$\mathcal{L}(\boldsymbol{\theta}\|\text{DATA})$	log-likelihood function for the unknown parameter vector $\boldsymbol{\theta}$
ρ_{ij}	conditional predictive distribution for system i failing due to component j
ρ_i	conditional predictive distribution for system i
N	total number of future failures
s	time on the calendar time scale after the DFD

I. Introduction

A. Background

Most products or systems consist of multiple components that are necessary for maintaining proper product functionality. For example, a laptop computer needs a fully functioning keyboard, monitor screen, CPU, and memory drive. If any one of these components were to fail, it would result in the failure of the entire system. A manufacturer is responsible for the repairs or replacements of components that fail during the warranty period. It is required that manufacturers maintain cash reserves to cover warranty costs for their sold products. More broadly, companies are often interested in the longer-term performance of their products, after the warranty period is over. In some applications, there are some more important concerns such as safety issues. Inaccurate predictions could lead to financial penalties or safety concerns. Thus it is always desirable to have accurate field-failure predictions.

Because the failure of components leads to product failure, the problem of modeling and predicting the failure time of such a product often involves a competing risk analysis. Classical competing risk models are often used in field-failure prediction for products with multiple failure modes (the failure of a particular component often corresponds to a particular failure mode). Most current prediction methods that use competing risks data assume that the behavior of each risk is constant over the production period of the product. Specifically, while the hazard function for a component may change over time, the hazard function used in the prediction model is treated to be the same for units manufactured over different production periods. It should be noted, however, that a product's design often evolves over a period of time with some components being replaced by alternative lower-cost or higher-reliability components, especially in the certain technological areas where product components are evolving rapidly. For example, a manufacturer may choose to make component replacements due to customer feedback, to improve their product's performance relative to a competitor, to incorporate new technological developments or for various other reasons regarding the performance of their product, or to reduce cost.

Many common household products' (e.g., televisions, laptop computers, and printers) system components are constantly being updated over time. Let us consider the laptop computer example given earlier. Even for the same laptop computer model from a specific manufacturer, different generations of hard drives may be used during the production period to improve the reliability or performance or to lower the cost of the laptop model. Figure 1 gives a graphical example of a product that has four major components, which we will call Product G. Over time, some of the component part numbers are changed while others stay

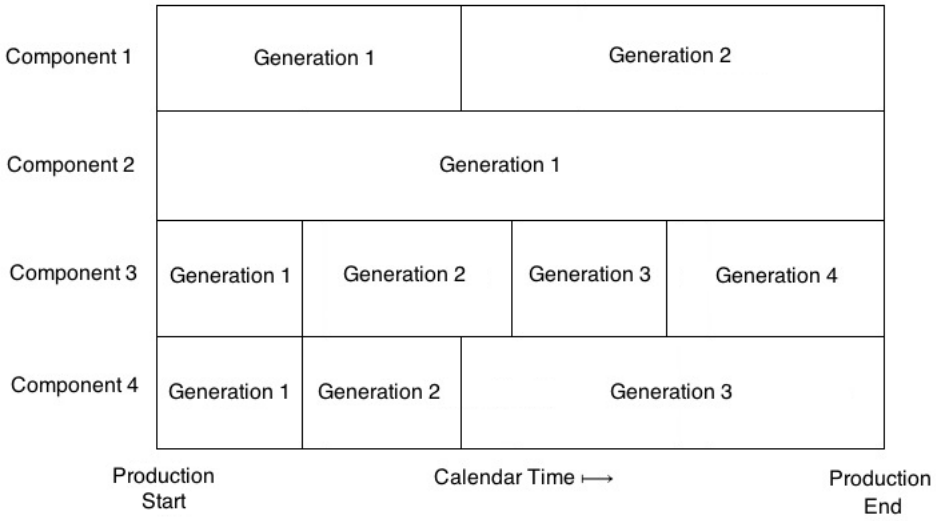

Figure 1: Example of product with multiple components of possibly multiple generations.

the same. The number of changes and the change times are generally not the same for all the components.

Products manufactured at different points in time will have different configurations, although consumers usually would not notice such changes. The overall reliability of the product, however, may change due to component part-number changes. Thus, the return rate of these products could depend on such changes. Product configuration information is generally available but is usually ignored in currently-used product-failure prediction models. Incorporating the product configuration information into field-failure prediction can provide important improvements over the classical prediction methods. A prediction model that does not consider product generation information is likely to be biased. One simple approach for handling the genealogy problem is to stratify the data according to the configuration groups and to perform separate analyses on each group. Doing so would, however, usually result in small amounts of data for some groups and is an inefficient use of available information. A better strategy is to "borrow strength" across different groups corresponding to particular part numbers by estimating common parameters. The major objective of this paper is to use generational information to make field-failure predictions. We will also compare different strategies and quantify the advantages of using generational information in predictions.

B. The Motivating Application

This research was motivated by a prediction problem that is similar to product D in Hong and Meeker [1]. Due to sensitive proprietary information, the actual data cannot be used here and a simulated dataset, similar to the original data, is used for illustration. The scenario, however, is close to the real application.

The dataset contains records of 6,000 system units, which entered into service at different times from March 2010 to March 2011, according to Figure 2. The failure of Product G is primarily caused by the failures of four key components. Based on early field returns, component four underwent several changes during production while component three was updated based on improvements in its manufacturing procedure. It was decided in mid 2010 to switch providers of component one for financial and quality reasons. Component two remained unchanged throughout the product's production. In March 2012, the dataset was frozen for analysis.

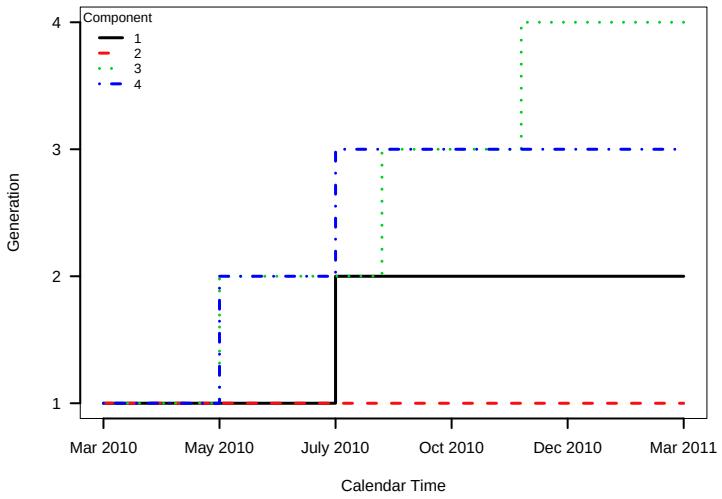

Figure 2: Generation functions for Product G.

C. Related Literature

A review of classical competing risks theory using latent response variables is given in Chiang [2], Moeschberger and David [3], Gail [4], David and Moeschberger [5], and Crowder [6]. Park and Kulasekera [7] use this approach to model and compare competing risks among different groups of samples from different populations. A second method of using a cause-specific hazard rate to model competing risks data is discussed in Prentice et.

al. [8]. A third approach using mixture models is explored in Larson and Dinse [9] and most recently in Maller and Zhou [10]. An excellent review of competing risks methods is given in Crowder [11].

A general introduction to warranty data analysis is given by Lawless [12] and Kalbfleisch, Lawless, and Robinson [13]. Methods for prediction in the presence of censored data are discussed in Escobar and Meeker [14]. Hong and Meeker [1] considered field-failure prediction with multiple failure modes. The most common tool used to quantify potential field-failure prediction error is the prediction interval (PI), with much of the literature focused on calibration methods for the naive "plug-in" procedure. A theoretical perspective is taken in Komaki [15], Barndorff-Nielson and Cox [16], and Vidoni [17], while Beran [18] and Escobar and Meeker [14] use simulation/resampling methods. The predictive distribution used in Lawless and Fredette [19] provides a useful method of presentation of the prediction interval procedure that is equivalent to calibration of the naive intervals. Hong and Meeker [1] adapted this approach for warranty-prediction applications involving multiple failure modes.

D. Overview

The rest of this paper is organized as follows. Section II describes the data and model used in this paper. Section III gives the maximum likelihood (ML) procedure for estimating the model parameters. Section IV gives the procedure for predicting the cumulative number of field failures at a future time for products in the risk set. Section V evaluates the advantage of incorporating component genealogy information relative to methods that ignore such information. Finally, Section VI presents a summary and discussion of areas for future research.

II. Data and Model

A. Product G Data

The failure-time data are denoted by $\{\tau_i, t_i, g_{ij}, \delta_{ij}; j = 1, \ldots, J\}$, $i = 1, 2, \ldots, n$. The value J denotes the number of components and n is the total number of observations. For the illustrative dataset, $J = 4$ and $n = 6{,}000$. Here, τ_i is the time of installation on the calendar time scale for system i and t_i is the observed time to failure (time in service) for the failed (censored) system i. The quantity δ_{ij} is the observed failure indicator for component j of system i. Here $\delta_{ij} = 1$ and $\delta_{ik} = 0$ $(k \neq j)$ if system i failed due to component j, and $\delta_{ij} = 0$ for all j if system i is censored. Finally, $g_{ij} = g(\tau_i, j)$ is the generation information function that gives the generation of component j based on the installation time τ_i for the

Table I: Failure proportion (the number of failures divided by the total number of system units), the number of failures, and number of system units surviving of Product G by component and generation at the DFD. Here "C", "Prop.", "Fail.", and "RS" stand for, component, failure proportion, number of failures, and the number of system units in the risk set, respectively. The total number of system units is 6,000.

C	Generation												Total	
	1			2			3			4				
	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.	RS	Prop.	Fail.
1	0.0055	33	1838	0.0038	23	3553		-			-		0.0093	56
2	0.0240	144	5391		-			-			-		0.0240	144
3	0.0048	29	777	0.0030	18	1520	0.0013	8	1426	0.0008	5	1668	0.0099	60
4	0.0528	317	777	0.0100	60	1061	0.0010	6	3553		-		0.0638	383

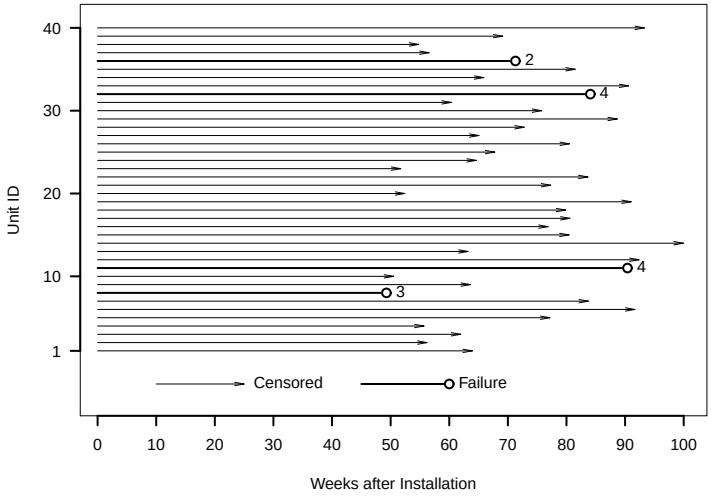

Figure 3: Event plot showing the lifetimes for the first 40 system units of Product G. The numbers next to the failure arrows indicate the component that caused failure.

system units. Let G_j be the maximum number of possible generations for component j. An example of the generation functions for Product G is presented in Figure 2. Of the 6,000 system units followed, 609 failed prior to the data freeze date (DFD). Table I shows the failure proportion (the number of failures divided by the total number of system units), the number of failures, and number of system units surviving of Product G by component and generation at the DFD.

B. Failure-time Model

The log-location-scale family of distributions includes the lognormal and Weibull distributions as special cases and will be used here to model the component lifetimes. Let T_{ij} denote the latent lifetime of component j of system i. Assuming the particular log-location-scale distribution does not change across generations, which is reasonable if the failure mode is the same within each component, the cumulative distribution function (cdf) of T_{ij} is given by

$$F_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) = \Phi_{ij}\left[\frac{\log(t) - \mu_{jg_{ij}}}{\sigma_j}\right] , \tag{1}$$

where $\mu_{jg_{ij}}$ and σ_j are the location and scale parameter respectively for $\log(T_{jg_{ij}})$, $\boldsymbol{\xi}_{jg_{ij}} = (\mu_{jg_{ij}}, \sigma_j)'$ is the parameter vector for component j of system i at the g_{ij}th generation, and $\Phi(\cdot)$ is the standard cdf for the log-location-scale family of distributions. Within the competing-risk model, F_{ij} is the marginal cdf for component j of system i. The probability density function (pdf) for $T_{jg_{ij}}$ is given by

$$f_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) = \frac{1}{\sigma_j t}\phi_{ij}\left[\frac{\log(t) - \mu_{jg_{ij}}}{\sigma_j}\right] , \tag{2}$$

where $\phi(\cdot)$ is the standard pdf for the log-location-scale family of distributions.

The primary focus of this paper will be on models in which only generational changes in the location parameter are considered, which we will refer to as "location-change" generational (LCG) models. However, this model can easily be extended to allow for generational changes in the scale parameter as well. We will refer to these models as "extended location-change" generational (ELCG) models. Because the scale parameter is often associated with the failure mechanism, the LCG model tends to be more appropriate with respect to most field applications. However, the ELCG model is more flexible and may be appropriate in cases where an updated component does contain a different failure mechanism than the previous generation. An example would be transitioning from a traditional hard drive to a solid-state hard drive in a laptop.

It is important to note that the location and scale parameters are used here in connection with their usage in the location-scale part of the log-location-scale distribution. For example, in the lognormal distribution, μ represents the location parameter and σ the scale parameter of the normal distribution, describing the logarithms of the times. This is to be distinguished from the shape and scale parameters of the lognormal, which can be represented by σ and $\exp(\mu)$, respectively. It is with respect to the former case in which we will be referring to the location and scale parameters of the distribution.

Many well known distributions used in lifetime modeling belong to the log-location-scale family. For illustration purposes, we will model the component lifetimes using the Weibull and lognormal distributions. The Weibull distribution cdf and pdf are given by replacing Φ and ϕ in (1) and (2) with the standard smallest extreme value distributions $\Phi_{\text{sev}}(z) = 1 - \exp\left[-\exp\left(z\right)\right]$, and $\phi_{\text{sev}}(z) = \exp\left[z - \exp\left(z\right)\right]$, respectively. The lognormal distribution cdf and pdf are obtained similarly, but with Φ_{nor} and ϕ_{nor}, which are, respectively, the cdf and pdf for the standard normal distribution.

C. System cdf and Sub-distribution Function

Under the classical competing risks framework, the distribution of a system's lifetime is the same as the distribution of the minimum latent lifetime over the J components. Because the components are independent of one another, the cdf of the system lifetime T_i is given by

$$F_i(t;\boldsymbol{\xi}_i) = \Pr[\min(T_{i1},\ldots,T_{T_{iJ}}) \leq t] = 1 - \prod_{j=1}^{J}\left[1 - F_{ij}(t;\boldsymbol{\xi}_{jg_{ij}})\right] \tag{3}$$

where $\boldsymbol{\xi}_i = (\boldsymbol{\xi}'_{1g_{i1}},\ldots,\boldsymbol{\xi}'_{Jg_{iJ}})'$ is the set of parameters for system i. An alternative expression for the system lifetime cdf is

$$F_i(t;\boldsymbol{\xi}_i) = \sum_{j=1}^{J} F^*_{ij}(t;\boldsymbol{\xi}_i). \tag{4}$$

Here, $F^*_{ij}(t;\boldsymbol{\xi}_i)$ is known as the sub-distribution function (e.g., see Moeschberger and David [3]), which is defined as

$$F^*_{ij}(t;\boldsymbol{\xi}_i) = \Pr(T_i \leq t, \Delta_{ij} = 1;\boldsymbol{\xi}_i), \tag{5}$$

where $\Delta_{ij} = 1$ and $\Delta_{il} = 0$ for all $l \neq j$ if component j is the cause of system failure. The $F^*_{ij}(t;\boldsymbol{\xi}_i)$ functions are needed to generate predictions for the individual failure modes. Peterson [20] showed that sub-distribution functions are not necessarily the same as true distribution functions as it may hold that $\lim_{t\to\infty} F^*_{ij}(t;\boldsymbol{\xi}_i) < 1$. The sub-distribution function is related to the marginal distributions of the components in classical competing risks theory by the following expression,

$$F^*_{ij}(t;\boldsymbol{\xi}_i) = \int_0^t f_{ij}(s;\boldsymbol{\xi}_{jg_{ij}}) \prod_{l\neq j}[1 - F_{il}(s;\boldsymbol{\xi}_{lg_{il}})]ds. \tag{6}$$

The sub-density function is

$$f^*_{ij}(t;\boldsymbol{\xi}_i) = \frac{dF^*_{ij}(t;\boldsymbol{\xi}_i)}{dt} = f_{ij}(t;\boldsymbol{\xi}_{jg_{ij}}) \prod_{l\neq j}[1 - F_{il}(t;\boldsymbol{\xi}_{lg_{il}})]. \tag{7}$$

III. Parameter Estimation

A. The Likelihood Function

The parameters of the generational model are estimated using ML estimation. In this section, we construct the likelihood function for the LCG and ELCG models. We will denote the set of distinct parameters to be estimated by $\boldsymbol{\theta}$.

Location-Change Generational Model: For the LCG model, the set of distinct parameters is denoted by $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_J')'$, where $\boldsymbol{\theta}_j = (\mu_{j1}, \ldots, \mu_{jG_j}, \sigma_j)'$. The likelihood is expressed as follow,

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{i=1}^{n} \left(\left\{ \prod_{j=1}^{J} f_{ij}^*(t_i; \boldsymbol{\xi}_i)^{\delta_{ij}} \right\} [1 - F_i(t_i; \boldsymbol{\xi}_i)]^{1-\sum_{j=1}^{J} \delta_{ij}} \right). \tag{8}$$

Note that $\boldsymbol{\xi}_i$ is defined in (3) as the set of parameters for system i, while $\boldsymbol{\theta}$ here is the set of distinct parameters in the LCG model. The construction of the likelihood function is based on the following facts. If system i failed due to component j, then the likelihood contribution is

$$\Pr(T_i = t_i; \Delta_{ij} = 1, \Delta_{il} = 0, l \neq j) = \Pr(T_{ij} = t_i; T_{il} > T_{ij}, l \neq j) = f_{ij}^*(t_i; \boldsymbol{\xi}_i).$$

If system i is censored, then the likelihood contribution is

$$\Pr(T_{ij} > t_i; j = 1, \ldots, J) = 1 - F_i(t_i; \boldsymbol{\xi}_i).$$

The likelihood function in (8) can be re-expressed as

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left\{ \prod_{i=1}^{n} f_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}. \tag{9}$$

The likelihood in (9) is equivalent to treating failures caused by components $l \neq j$ as if they were censored for component j, which is a common practice in competing risks analysis and is also appropriate here because generational information is specific to a component and is not affected by other components under our assumptions. One can also maximize the likelihood in (9) separately for each component because there are no common parameters across components.

Extended Location-Change Generational Model: For the ELCG model, the set of distinct parameters is denoted by $\boldsymbol{\theta} = (\boldsymbol{\theta}_{11}^{*'}, \ldots, \boldsymbol{\theta}_{1G_1}^{*'}, \ldots, \boldsymbol{\theta}_{J1}^{*'}, \ldots, \boldsymbol{\theta}_{JG_J}^{*'})'$. Here $\boldsymbol{\theta}_{jg}^* = (\mu_{jg}, \sigma_{jg})'$, and μ_{jg} and σ_{jg} are location and scale parameters for the log lifetime of the gth generation of component j, respectively. The likelihood can be expressed as

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left(\prod_{g=1}^{G_j} \left\{ \prod_{i \in \text{DATA}_{jg}} f_{ij}(t_i; \boldsymbol{\theta}_{jg}^*)^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\theta}_{jg}^*)]^{1-\delta_{ij}} \right\} \right). \tag{10}$$

where DATA_{jg} is the portion of the sample that has component j belonging to generation g. Here, the focus is on the lifetime contribution for a particular generation of the component failures due to other generations within the same component or in other components being treated as censored observations.

B. Maximum Likelihood Estimates and Information Matrix

The ML estimate, denoted by $\widehat{\boldsymbol{\theta}}$, is obtained by maximizing (9) for the LCG model or (10) for the ELCG model. Depending on the distribution of each component, the solution may have a closed-form expression, but more often a numerical optimization procedure is needed. For our analysis, the ML estimates were calculated using the BFGS quasi-Newton iterative procedure developed by Broyden [21], Fletcher [22], Goldfarb [23], and Shanno [24]. When estimating the scale parameter, a log transformation was used so that the optimization was unbounded. Also, when there is heavy censoring, it is useful to replace μ by a small quantile, so that the likelihood is more well-behaved.

The local information matrix is obtained by evaluating the negative Hessian at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. In particular, the local information matrix is

$$I(\widehat{\boldsymbol{\theta}}) = -\frac{\partial^2 \log[L(\boldsymbol{\theta}|\text{DATA})]}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}.$$

The details of the calculation of the local information matrix for each model are given in Appendix A. An estimate of the variance-covariance matrix of the parameter estimators, $\widehat{\Sigma}$, is obtained by evaluating $[I(\boldsymbol{\theta})]^{-1}$ at $\widehat{\boldsymbol{\theta}}$.

Wald confidence intervals (CIs) can be calculated for the ML estimates using $\widehat{\Sigma}$. As an example, a $100(1-\alpha)\%$ Wald CI for μ_{jg} in the LCG model is given by

$$\widehat{\mu}_{jg} \pm z_{1-\alpha/2}\sqrt{\widehat{\text{var}}(\widehat{\mu}_{jg})} \tag{11}$$

where $z_{1-\alpha/2}$ is the $100(1-\alpha/2)\%$ quantile of the standard normal distribution and $\widehat{\text{var}}(\widehat{\mu}_{jg})$ is the estimated variance of $\widehat{\mu}_{jg}$ taken from the corresponding diagonal element of $\widehat{\Sigma}$.

C. Estimation for Product G Data

We apply the ML method for the LCG model to the Product G data to estimate the unknown model parameters. Table II gives the ML estimates, standard errors, and the approximate 95% CIs for the model parameters for each component and each generation.

Table II: Summary of ML estimation for Product G data. The time unit is weeks.

Comp.	Gen.	Param.	True Values	Estimate	Std. Error	95% Lower	95% Upper
1	1	μ	6.20	6.080	0.213	5.662	6.498
	2	μ	6.30	6.055	0.250	5.564	6.545
	All	σ	0.40	0.369	0.129	0.286	0.475
2	1	μ	5.00	4.991	0.035	4.923	5.059
		σ	0.30	0.297	0.060	0.264	0.334
3	1	μ	5.63	5.635	0.148	5.345	5.925
	2	μ	5.73	5.739	0.181	5.384	6.094
	3	μ	5.83	5.768	0.219	5.339	6.196
	4	μ	5.93	5.747	0.251	5.254	6.240
	All	σ	0.30	0.291	0.127	0.227	0.374
4	1	μ	4.68	4.670	0.010	4.651	4.690
	2	μ	4.78	4.751	0.018	4.716	4.785
	3	μ	4.88	4.810	0.033	4.746	4.875
	All	σ	0.20	0.186	0.042	0.171	0.202

IV. Field-Failure Predictions

A. Point Predictions and Intervals

For prediction, the first step is to find point predictions. Then, for most applications, the next step is to find PIs on the point predictions. Both tasks can be accomplished using conditional predictive distributions (Lawless and Fredette [19]), which are probability distributions for the remaining lifetime of a system given its current lifetime. These distributions are presented here in terms of the LCG model. The computing of the predictive distributions for the ELCG model are similar.

The conditional predictive distribution for failure mode j of system i is

$$\rho_{ij} = \rho_{ij}(s) = \Pr(\Delta_{ij} = 1, T_i \leq t_i + s | T_i > t_i) = \frac{F^*_{ij}(t_i + s; \boldsymbol{\xi}_i) - F^*_{ij}(t_i; \boldsymbol{\xi}_i)}{1 - F_i(t_i; \boldsymbol{\xi}_i)}, \tag{12}$$

for some $s > 0$. The ML estimate of ρ_{ij}, denoted by $\widehat{\rho}_{ij}$, can be obtained by evaluating (12) at $\widehat{\boldsymbol{\theta}}$. The conditional predictive distribution for system i is

$$\rho_i = \rho_i(s) = \Pr(T_i \leq t_i + s | T > t_i) = \frac{F_i(t_i + s; \boldsymbol{\xi}_i) - F_i(t_i; \boldsymbol{\xi}_i)}{1 - F_i(t_i; \boldsymbol{\xi}_i)}, \tag{13}$$

for some $s > 0$. The ML estimator of ρ_i, denoted by $\widehat{\rho}_i$, can be obtained by evaluating (13) at $\widehat{\boldsymbol{\theta}}$. Note that here $\rho_i = \sum_{j=1}^{J} \rho_{ij}$. The point prediction for the total number of failures is given by $\widehat{N} = \widehat{N}(s) = \sum_{i \in \text{RS}} \widehat{\rho}_i$, where RS is the risk set containing systems that have not yet failed by the DFD.

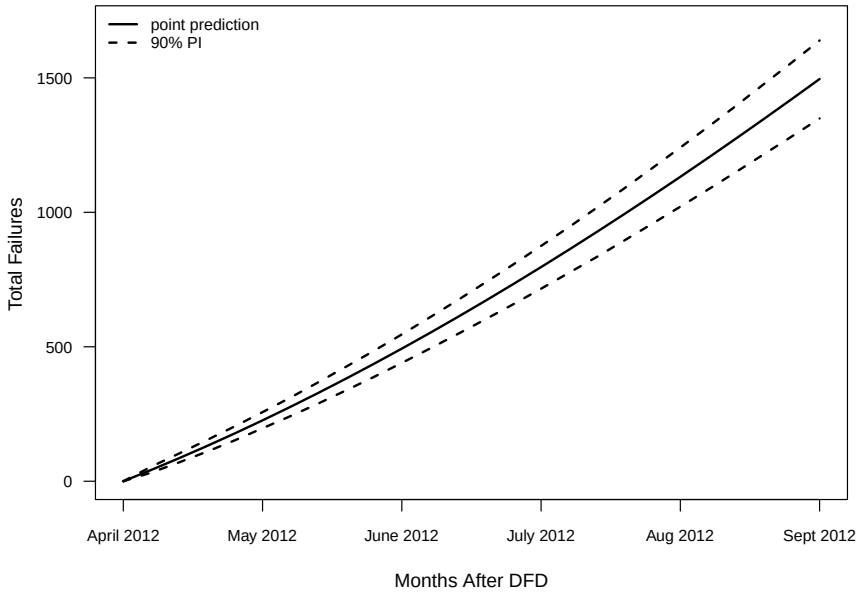

Figure 4: The predicted cumulative number of failures as function of time after the DFD, plotted in the calendar time scale.

B. Prediction Results for Product G

Point predictions can be computed as a function of time after the DFD and PIs can also be generated for each point prediction. In particular, the point prediction is obtained as $\widehat{N} = \sum_{i \in \text{RS}} \widehat{\rho}_i$ at the ML estimate. The computing of PIs is based on an extension of the procedure given in Hong and Meeker [1]. The details of the calculations are discussed in Appendix B. Figure 4 shows the point predictions and PIs for the cumulative number of failures for the at-risk system units as a function of time after the DFD.

V. Assessment of Model by Simulation

A. Prediction Models

In Section II-B, we introduced the LCG and ELCG models. In this section, we will compare the standard competing risks (SCR) model, the LCG model, and the ELCG model. The SCR model corresponds to the commonly-used prediction procedure in which no generational changes are considered and all data are pooled to estimate a failure distribution for each

component. Thus, the likelihood for the SCR model simplifies to

$$L(\boldsymbol{\theta}|\text{DATA}) = \prod_{j=1}^{J} \left\{ \prod_{i=1}^{n} f_{ij}(t_i; \boldsymbol{\theta}_j^*)^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\theta}_j^*)]^{1-\delta_{ij}} \right\}, \tag{14}$$

where t_i is the lifetime of system i, $\boldsymbol{\theta}_j^* = (\mu_j, \sigma_j)'$ is the vector of location and scale parameters for the log of the lifetime of component j. The set of distinct parameters for the SCR model is $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^{*\prime}, \ldots, \boldsymbol{\theta}_J^{*\prime})'$. The derivation of the local information matrix is given in Appendix A. The conditional predictive distributions used in the point prediction and PI calculation are similar to those given in (12) and (13).

B. Comparison Criterion

The assessment of each model is based on the accuracy of its prediction of the cumulative total number of failures at a specified time point after the DFD. For each simulated dataset, the mean squared error (MSE) of its prediction will be estimated as a function of future time. The estimated MSE is calculated as follows:

$$\widehat{\text{MSE}} = \frac{1}{B} \sum_{b=1}^{B} (\widehat{N}_b - \widehat{N})^2, \tag{15}$$

where $\widehat{N}_b$ is the estimated prediction based on bootstrap simulation b, B is the total number of bootstrap samples, and $\widehat{N} = \sum_{i \in \text{RS}} \widehat{\rho}_i$ is the point prediction based on the simulated dataset. The bootstrap samples were generated based on the random-weight bootstrapping procedure. Details on this procedure are given in Appendix B. Here, we use bootstrapping to estimate the MSE for each dataset because each simulated dataset will have a different size risk set. Thus, the models were compared based on the average relative MSE (RMSE) of the MSEs calculated for each dataset at individual time points after the DFD using the LCG model as a basis. For example, the RMSE for SCR vs LCG is computed as the ratio of the MSE of prediction based on the SCR model and the MSE of prediction based on the LCG model for each dataset. A large number of datasets (e.g., 200) were simulated in order to ensure a stable average of the RMSE.

C. Simulation Setup

To investigate performance of our models over a wide range of situations, our simulation considers eight scenarios. In each scenario, lifetime data were simulated based on the competing risks model, in which each risk has a specified lifetime distribution and the overall lifetime of a system is determined by the minimum lifetime of the independent competing

risks. Each dataset consisted of 6,000 system units with four competing risks, much like the original data. The lifetimes of the components were randomly generated using staggered entry and the distribution form remained the same across generations with the parameters of the distribution being dependent on the date the product was manufactured. After a set period of time, no more lifetimes were generated, corresponding to the end of "production." The values of the parameters were chosen to make the total proportion failing similar to the original data.

Table III summarizes the generation scenarios, distribution information, and parameter values used in the simulation. In the simulation, we use the LCG model as the true model. That is, the generation changes only occur to the location parameter μ and the scale parameter σ remains constant over generations. To aid in understanding, we present the changes in location parameter in terms of changes in the 0.1 quantile of the component lifetime distribution. That is, we use the percentage $100 \times \exp(\mu + \Delta\mu + z_{0.1}\sigma)/\exp(\mu + z_{0.1}\sigma)\%$, where $\Delta\mu$ is the amount of change and the $z_{0.1}$ is the 0.1 quantile of the standard location-scale distribution. The percentage changes presented in Table III are rounded to the nearest hundredth.

For the first five scenarios, the change across generations (if it exists) is the same for all generations where the amount of change is specific to each scenario. This change in $t_{0.1}$ ranges from no change (a standard competing risks model) to an increase of 65%. The first five scenarios will be used to test the sensitivity of the LCG and ELCG models to the magnitude of the generational change.

The remaining three scenarios consider more general cases of changes in $t_{0.1}$. The sixth scenario allows the magnitude of changes to vary across components but remaining the same within each component. The seventh scenario consists of changes in $t_{0.1}$ that gradually increase over time. Note that the change for component 3 in this scenario is represented by (5)(11)(28) in Table III, which means the consecutive changes are 5%, 11%, and 28%, respectively. The last scenario implements a decrease in lifetime across two selected generations.

D. Simulation Results

For each scenario, 200 datasets were simulated with estimation and point prediction performed for each dataset. Figure 5 gives the results for each scenario. The SCR and LCG models are nearly indistinguishable in the first scenario, which indicates that the LCG model is still adequate for situations in which no generational changes are present. In the remaining

Table III: Summary of the simulation scenarios.

Component	Dist.	μ	σ	# Gens.	% Change in $t_{0.1}$ for eight different scenarios (S1-S8)							
					S1	S2	S3	S4	S5	S6	S7	S8
1	Weibull	6.20	0.40	2	0	5	11	28	65	5	5	11
2	Lognormal	5.00	0.30	1	-	-	-	-	-	-	-	-
3	Weibull	5.63	0.30	4	0	5	11	28	65	11	(5)(11)(28)	(11)(-28)(11)
4	Lognormal	4.68	0.20	3	0	5	11	28	65	28	(5)(11)	(11)(-5)

scenarios, the SCR model performs worse than the LCG when generational changes are introduced. This is due to an increase in both the variance and bias of the SCR model predictions. It is a clear indication of the inflexibility of modeling this type of data with the standard approach. The tapering effect in all of the plots is due to the number of system units in the risk sets decreasing as time goes on (eventually all system units will fail), meaning predictions are slowly becoming more stable. It is interesting to note that the behavior of the SCR model in Scenario 8 is not as poor as in the others. This is most likely due to the large decrease in the quantile $t_{0.1}$ compared to the increases, which yields data that is more similar to that generated by a situation with no generational changes.

It is surprising to see the ELCG model performing more erratically than the more restricted LCG. Further simulations were performed using progressively larger sample sizes while maintaining a consistent fraction failing to determine if the number of expected failures might be the cause. In the computing of the variance and bias, the normal approximation was used to simulate data, instead of using bootstrap, to reduce the time needed for simulations. The large size (e.g., 50,000) of several of these new simulations prevents us doing the bootstrap in a reasonable time frame. Figure 6 gives the variance and squared bias as a function of time after the DFD for sample size $n = 6{,}000, 10{,}000,$ and $50{,}000$. Both the variance and squared bias of the prediction were divided by the sample size n so that a fair comparison can be made across different sample sizes. Figure 6 reveals the prediction variance to be the dominant factor in the MSE. As the sample increases, the variance of the early predictions decreases as does the overall bias, confirming the sensitivity of the ELCG model to the number of failures. This is because estimation of the parameters in the ELCG model is performed at the generational level and so more failures are needed in every generation, especially the later ones, to yield more accurate predictions. At a sample size of $n = 6{,}000$, Table I shows the observed fraction failing for the fourth generation of component 3 is only 0.08% (5 failures), providing further insight into why the variance is large for the ELCG model. This simulation also shows an advantage of the LCG model in its ability to borrow

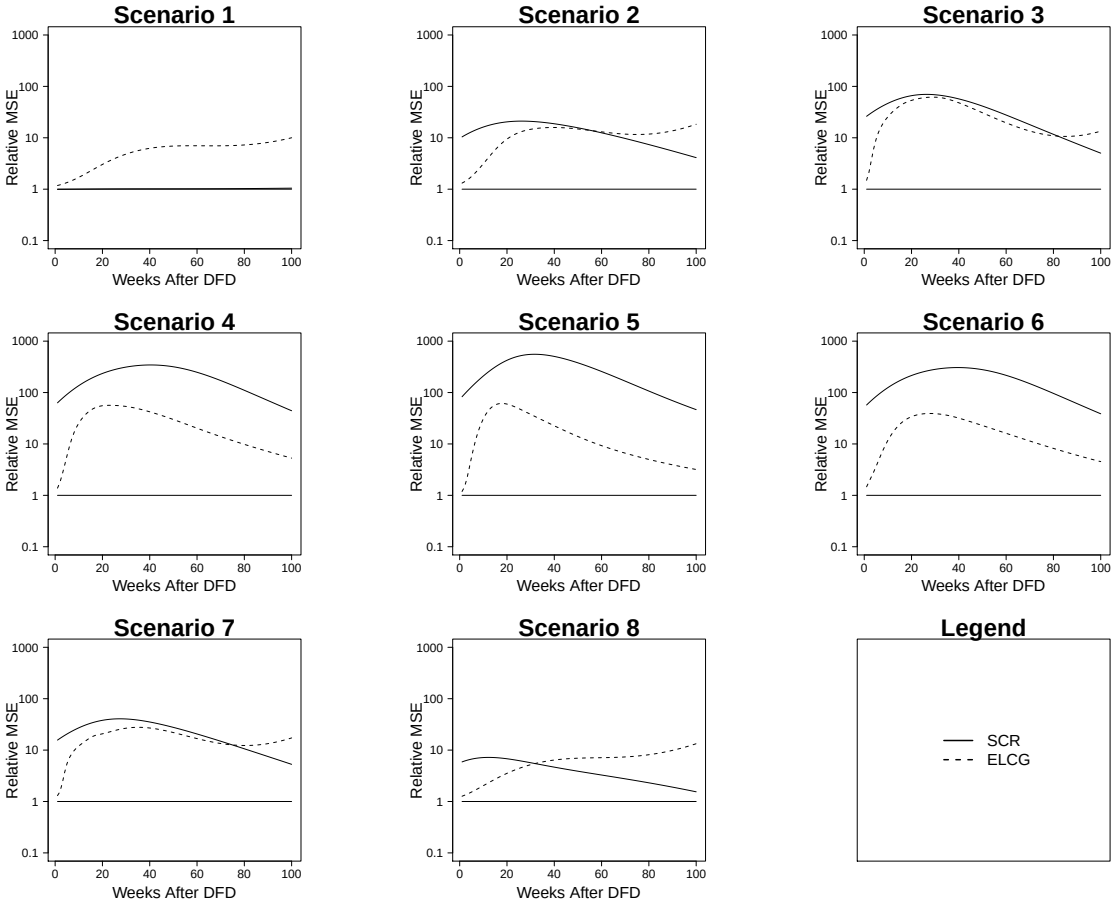

Figure 5: Comparison of SCR and ELCG models relative to the LCG model via relative MSE (RMSE) for each scenario. The horizontal line corresponds to RMSE$=1$ for the LCG model.

strength across generations.

VI. Conclusions and Areas for Future Research

This paper presented an extension of the standard competing risks model to include the possibility of risks changing over the time that a product was manufactured. Methods for estimating parameters and calculating conditional predictive distributions were proposed. An assessment of the generational model and an extended generational model was performed using cumulative predictions under eight different scenarios. It was found that the SCR model performed poorly relative to the LCG model when generational changes were present in the data.

The approach to field-failure predictions presented in this paper is general. Specific applications, such as warranty return prediction can be easily accommodated by the method

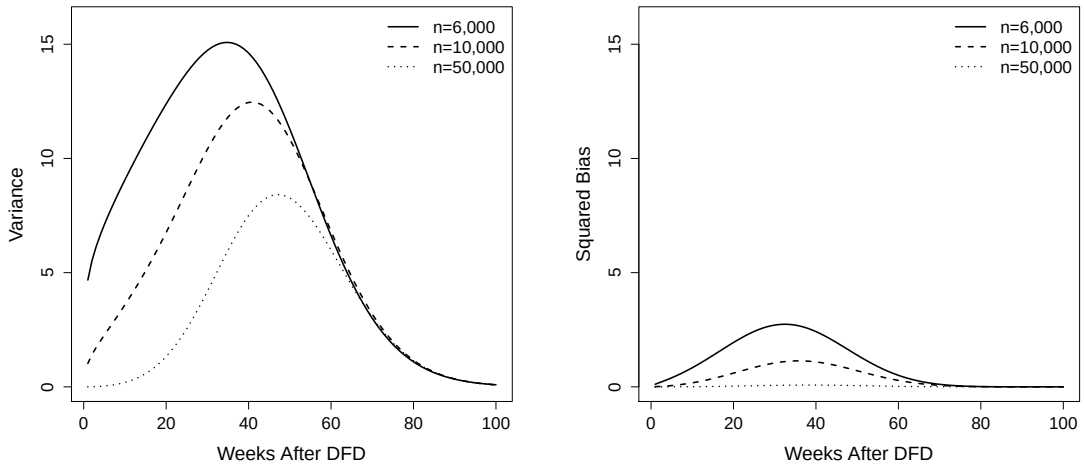

Figure 6: Variance and squared bias of ELCG model predictions adjusted for sample size.

in this paper. One only needs to adjust the risk accordingly for those system units with expired warranty (e.g., typically one-year to three-year period). There are also situations that a company also needs to predict warranty costs not just for system units that have already gone to the field, but also for system units that will be produced put into the field in the future. One only needs to add future system units to the risk set and the predictions can be obtained correspondingly.

The models presented here were basic in their construction and several restrictions were imposed in order to obtain a suitable foundation for the idea. In future work, these restrictions can be removed so as to formulate more complicated models. The models considered here assumed consistent changes in the model parameters. Areas for future research could include random changes in the model parameters or perhaps even in the distribution form. Allowing for covariates, especially dynamic covariates, in a regression setting could also be considered. Methods for lifetime predictions using dynamic covariates have been discussed in Hong and Meeker [25] and similar procedures could be adapted here. Heterogeneities in operating environments (e.g., Ye et al. [26]) can also be considered with generational changes.

In this model, the components were assumed to be independent of one another. This assumption, however, is not always true for products in the field. Although components may be grouped together to allow for independence among groups, incorporating dependence among the components and investigating how it varies across generations and affects predictions may be useful in some applications.

ACKNOWLEDGMENTS

The work by King and Hong was partially supported by the National Science Foundation under Grant CMMI-1068933 to Virginia Tech and the 2011 DuPont Young Professor Grant.

APPENDIX A
LIKELIHOOD INFORMATION MATRIX COMPUTATION

A. Location-Change Generational Model

The likelihood given in (9) can be divided into a product of sub-likelihoods at the component level. These sub-likelihoods are given by

$$L^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}}[1 - F_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}, \tag{16}$$

where DATA_j is the portion of the sample associated with component j. These sub-likelihoods are useful in determining the local information matrix of the parameters.

The log-likelihood $\mathcal{L}(\boldsymbol{\theta}|\text{DATA}) = \log[L(\boldsymbol{\theta}|\text{DATA})]$ is the sum of the logarithms of the sub-likelihoods, which are given as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \sum_{i=1}^{n} \left\{ \delta_{ij} \log[f_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})] + (1-\delta_{ij}) \log[1 - F_{ij}(t_i;\boldsymbol{\xi}_{ig_{ij}})] \right\}, \tag{17}$$

Because no system unit can belong to two different generations within the same component, (17) can be re-expressed as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j|\text{DATA}_j) = \sum_{g=1}^{G_j} \sum_{i \in \text{DATA}_{jg}} \left\{ \delta_{ij} \log[f_{ij}(t_i;\boldsymbol{\xi}_{ig})] + (1-\delta_{ij}) \log[1 - F_{ij}(t_i;\boldsymbol{\xi}_{ig})] \right\}, \tag{18}$$

where the subscript g indicates the generation and DATA_{jg} is the portion of the sample associated with generation g of component j.

Because of the log-likelihood can be expressed as the sum of sub-log-likelihoods, it follows that

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{jg'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_j \partial \sigma_{j'}} = 0$$

for all $j \neq j'$ and $g \neq g'$. The local information matrix is then a $(G+J) \times (G+J)$

block-diagonal matrix, where $G = \sum_{j=1}^{J} G_j$, with J sub-information matrices given by

$$I^{(j)} = \begin{bmatrix} I_{\mu_{j1}} & 0 & \cdots & 0 & I_{\mu_{j1}\sigma_j} \\ 0 & I_{\mu_{j2}} & \cdots & 0 & I_{\mu_{j2}\sigma_j} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_{\mu_{jG_j}} & I_{\mu_{jG_j}\sigma_j} \\ I_{\mu_{j1}\sigma_j} & I_{\mu_{j2}\sigma_j} & \cdots & I_{\mu_{jG_j}\sigma_j} & I_{\sigma_j} \end{bmatrix}.$$

Using the notation given in (1) and (2) for log-location-scale distributions, we define the following functions for arbitrary μ and σ:

$$r_{ij}(z) = \frac{\phi'_{ij}(z)}{\phi_{ij}(z)}, \text{ and } h_{ij}(z) = \frac{\phi_{ij}(z)}{1 - \Phi_{ij}(z)}.$$

Here $z = [\log(t) - \mu]/\sigma$ and $\phi'_{ij}(z)$ is the first derivative of $\phi_{ij}(z)$ with respect to z. Using these functions, it follows that

$$I_{\mu_{jg}} = \frac{1}{\sigma_j^2} \sum_{i \in \text{DATA}_{jg}} \left[\delta_{ij} r'_{ij}(z_{ijg}) - (1 - \delta_{ij}) h'_{ij}(z_{ijg})\right]$$

$$I_{\sigma_j} = \frac{1}{\sigma_j^2} \sum_{g=1}^{G_j} \Bigg(\sum_{i \in \text{DATA}_{jg}} \left\{\delta_{ij} \left[z_{ijg}^2 r'_{ij}(z_{ijg}) + 2 z_{ijg} r_{ij}(z_{ijg}) + 1\right] \right.$$
$$\left. -(1 - \delta_{ij}) \left[z_{ijg}^2 h'_{ij}(z_{ijg}) + 2 z_{ijg} h_{ij}(z_{ijg})\right]\right\} \Bigg)$$

$$I_{\mu_{jg}\sigma_j} = \frac{1}{\sigma_j^2} \sum_{i \in \text{DATA}_{jg}} \left\{\delta_{ij} \left[z_{ijg} r'_{ij}(z_{ijg}) + r_{ij}(z_{ijg})\right] - (1 - \delta_{ij}) \left[z_{ijg} h'_{ij}(z_{ijg}) + h_{ij}(z_{ijg})\right]\right\}$$

where $z_{ijg} = [\log(t_i) - \mu_{jg}]/\sigma_j$.

B. Standard Competing Risk Model

In a similar manner for the LCG model, the SCR likelihood given in (14) can also be written as a product of J component-level sub-likelihoods given by

$$L^{(j)}(\boldsymbol{\theta}_j^* | \text{DATA}_j) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i; \boldsymbol{\theta}_j^*)^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\theta}_j^*)]^{1 - \delta_{ij}} \right\}$$

with the corresponding log-sub-likelihoods given as

$$\mathcal{L}^{(j)}(\boldsymbol{\theta}_j^* | \text{DATA}_j) = \sum_{i=1}^{n} \left\{ \delta_{ij} \log\left[f_{ij}(t_i; \boldsymbol{\theta}_j^*)\right] + (1 - \delta_{ij}) \log\left[1 - F_{ij}(t_i; \boldsymbol{\theta}_j^*)\right] \right\}.$$

Once again, it follows that

$$\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_j \partial \mu_{j'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_j \partial \sigma_{j'}} = 0$$

for all $j \neq j'$, where $\mathcal{L}(\boldsymbol{\theta}|\text{DATA})$ is the logarithm of (14). The local information matrix is a $(2J) \times (2J)$ block-diagonal matrix of J sub-information matrices given by

$$I^{(j)} = \begin{bmatrix} I_{\mu_j} & I_{\mu_j \sigma_j} \\ I_{\mu_j \sigma_j} & I^*_{\sigma_j} \end{bmatrix}$$

where

$$\begin{aligned}
I_{\mu_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left[\delta_{ij} r'_{ij}(z_{ij}) - (1 - \delta_{ij}) h'_{ij}(z_{ij}) \right] \\
I^*_{\sigma_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left\{ \delta_{ij} \left[z_{ij}^2 r'_{ij}(z_{ij}) + 2 z_{ij} r_{ij}(z_{ij}) + 1 \right] - (1 - \delta_{ij}) \left[z_{ij}^2 h'_{ij}(z_{ij}) + 2 z_{ij} h_{ij}(z_{ij}) \right] \right\} \\
I_{\mu_j \sigma_j} &= \frac{1}{\sigma_j^2} \sum_{i=1}^{n} \left\{ \delta_{ij} \left[z_{ij} r'_{ij}(z_{ij}) + r_{ij}(z_{ij}) \right] - (1 - \delta_{ij}) \left[z_{ij} h'_{ij}(z_{ij}) + h_{ij}(z_{ij}) \right] \right\}.
\end{aligned}$$

Here, $z_{ij} = [\log(t_i) - \mu_j]/\sigma_j$.

C. Extended Location-Change Generational Model

For the ELCG model, the likelihood given in (10) can be divided at the generation level with the sub-likelihoods given as

$$L^{(jg)}(\boldsymbol{\theta}^*_{jg}|\text{DATA}_{jg}) = \prod_{i \in \text{DATA}_{jg}} f_{ij}(t_i; \boldsymbol{\xi}_{ig})^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\xi}_{ig})]^{1-\delta_{ij}}. \tag{19}$$

and the corresponding sub-log-likelihoods as

$$\mathcal{L}^{(jg)}(\boldsymbol{\theta}^*_{jg}|\text{DATA}_{jg}) = \sum_{i \in \text{DATA}_{jg}} \delta_{ij} \log [f_{ij}(t_i; \boldsymbol{\xi}_{ig})] + (1 - \delta_{ij}) \log [1 - F_{ij}(t_i; \boldsymbol{\xi}_{ig})]. \tag{20}$$

As with the previous models, it follows that

$$\begin{aligned}
&\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{jg'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \mu_{jg} \partial \mu_{j'g'}} \\
&= \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_{jg} \partial \sigma_{j'g}} = \frac{\partial^2 \mathcal{L}}{\partial \sigma_{jg} \partial \sigma_{jg'}} = \frac{\partial^2 \mathcal{L}(\boldsymbol{\theta}|\text{DATA})}{\partial \sigma_{jg} \partial \sigma_{j'g'}} = 0
\end{aligned}$$

for all $j \neq j'$ and $g \neq g'$, where $\mathcal{L}(\boldsymbol{\theta}|\text{DATA})$ is the logarithm of (10). Thus, the local information matrix is then a $(2G) \times (2G)$ block-diagonal matrix with G sub-information matrices given by

$$I^{(jg)} = \begin{bmatrix} I^*_{\mu_{jg}} & I_{\mu_{jg} \sigma_{jg}} \\ I_{\mu_{jg} \sigma_{jg}} & I_{\sigma_{jg}} \end{bmatrix}$$

where

$$I^*_{\mu_{jg}} = \frac{1}{\sigma^2_{jg}} \sum_{i \in \text{DATA}_{jg}} \left[\delta_{ij} r'_{ij}(z^*_{ijg}) - (1-\delta_{ij}) h'_{ij}(z^*_{ijg})\right]$$

$$I_{\sigma_{jg}} = \frac{1}{\sigma^2_{jg}} \sum_{i \in \text{DATA}_{jg}} \left\{\delta_{ij}\left[z^{*2}_{ijg} r'_{ij}(z^*_{ijg}) + 2 z^*_{ijg} r_{ij}(z^*_{ijg}) + 1\right]\right.$$
$$\left. -(1-\delta_{ij})\left[z^{*2}_{ijg} h'_{ij}(z^*_{ijg}) + 2 z^*_{ijg} h_{ij}(z^*_{ijg})\right]\right\}$$

$$I_{\mu_{jg}\sigma_{jg}} = \frac{1}{\sigma^2_{jg}} \sum_{i \in \text{DATA}_{jg}} \left\{\delta_{ij}\left[z^*_{ijg} r'_{ij}(z^*_{ijg}) + r_{ij}(z^*_{ijg})\right] - (1-\delta_{ij})\left[z^*_{ijg} h'_{ij}(z^*_{ijg}) + h_{ij}(z^*_{ijg})\right]\right\}.$$

Here, $z^*_{ijg} = [\log(t_i) - \mu_{jg}]/\sigma_{jg}$.

APPENDIX B
PREDICTION PROCEDURE

Here we briefly describe the procedure for obtaining PIs in the context of the LCG model. The procedure is similar to Hong and Meeker [1] but is extended for the LCG model.

A. Sampling Distribution

1) Simulate random values $Z_i, i = 1, \ldots, n$ that are independent and identically distribution with a distribution, which has the property that $\text{E}(Z_i) = \sqrt{\text{Var}(Z_i)}$. For our simulation, the distribution of choice was the exponential distribution with $\text{E}(Z_i) = 1$.
2) The weighted likelihood is computed as

$$L^*(\boldsymbol{\theta}|\text{DATA}) = \prod_{i=1}^{n} [L_i(\boldsymbol{\xi}_i|\text{DATA})]^{Z_i}$$

where

$$L_i(\boldsymbol{\xi}_i|\text{DATA}) = \prod_{i=1}^{n} \left\{ f_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})^{\delta_{ij}} [1 - F_{ij}(t_i; \boldsymbol{\xi}_{ig_{ij}})]^{1-\delta_{ij}} \right\}^{Z_i}.$$

3) Obtain the ML estimate $\widehat{\boldsymbol{\theta}}^*$ by maximizing $L^*(\boldsymbol{\theta}|\text{DATA})$.
4) Repeat steps 1-3 B times to get B bootstrap samples $\widehat{\boldsymbol{\theta}}^*_b, b = 1, 2, \ldots, B$.

B. Prediction Intervals

1) Simulate I^*_i from Bernoulli$(\widehat{\rho}_i), i \in$ RS, and compute $N^* = \sum_{i \in RS} I^*_i$.
2) Repeat step 1 B times to get $N^*_b, b = 1, 2, \ldots, B$.
3) Obtain $\widehat{\boldsymbol{\theta}}^*_b$ by using the procedure described in B.1

4) Compute $U_b^* = F_N(N_b^*; \widehat{\boldsymbol{\theta}}_b^*), b = 1, 2, \ldots, B$, where $F_N(\cdot)$ is the cdf of the Poisson-Binomial distribution with probability vector $\widehat{\rho}_i$'s, as given in Hong [27].
5) Let u_N^l, u_N^u be, respectively, the lower and upper $\alpha/2$ sample quantiles of $U_b^*, b = 1, 2, \ldots, B$. Compute the $100(1-\alpha)\%$ calibrated PI by solving for N_l and N_u in $F_N(N_l; \widehat{\boldsymbol{\theta}}) = u_N^l$ and $F_N(N_u; \widehat{\boldsymbol{\theta}}) = u_N^u$, respectively.

REFERENCES

[1] Y. Hong and W. Q. Meeker, "Field-failure and warranty prediction based on auxiliary use-rate information," *Technometrics*, vol. 52, no. 2, pp. 148–159, 2010.

[2] C. L. Chiang, "A stochastic study of the life table and its applications. III. The follow-up study with the consideration of competing risks," *Biometrics*, vol. 17, no. 1, pp. 57–78, 1961.

[3] M. L. Moeschberger and H. A. David, "Life tests under competing causes of failure and the theory of competing risks," *Biometrics*, vol. 27, no. 4, pp. 909–933, 1971.

[4] M. Gail, "A review and critique of some models used in competing risk analysis," *Biometrics*, vol. 31, no. 1, pp. 209–222, 1975.

[5] H. A. David and M. L. Moeschberger, *The Theory of Competing Risks*. Griffin, 1978.

[6] M. J. Crowder, *Classical Competing Risks*. CRC Press, 2001.

[7] C. Park and K. Kulasekera, "Parametric inference on incomplete data with competing risks among several groups," *IEEE Transactions on Reliability*, vol. 53, no. 1, pp. 11–21, 2004.

[8] R. L. Prentice, J. D. Kalbfleisch, A. V. Peterson, N. Flournoy, V. T. Farewell, and N. E. Breslow, "The analysis of failure times in the presence of competing risks," *Biometrics*, vol. 34, no. 4, pp. 541–554, 1978.

[9] M. G. Larson and G. E. Dinse, "A mixture model for the regression analysis of competing risks data," *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 34, no. 3, pp. 201–211, 1985.

[10] R. A. Maller and X. Zhou, "Analysis of parametric models for competing risks," *Statistica Sinica*, vol. 12, no. 3, pp. 725–750, 2002.

[11] M. J. Crowder, *Multivariate Survival Analysis Competing Risks*. CRC Press, 2012.

[12] J. F. Lawless, "Statistical analysis of product warranty data," *International Statistical Review / Revue Internationale de Statistique*, vol. 66, no. 1, pp. 41–60, 1998.

[13] J. D. Kalbfleisch, J. F. Lawless, and J. A. Robinson, "Methods for the analysis and prediction of warranty claims," *Technometrics*, vol. 33, no. 3, pp. 273–285, 1991.

[14] L. A. Escobar and W. Q. Meeker, "Statistical prediction based on censored life data," *Technometrics*, vol. 41, no. 2, pp. 113–124, 1999.

[15] F. Komaki, "On asymptotic properties of predictive distributions," *Biometrika*, vol. 83, no. 2, pp. 299–313, 1996.

[16] O. E. Barndorff-Nielsen and D. R. Cox, "Prediction and asymptotics," *Bernoulli*, vol. 2, no. 4, pp. 319–340, 1996.

[17] P. Vidoni, "Improved prediction intervals and distribution functions," *Scandinavian Journal of Statistics*, vol. 36, no. 4, pp. 735–748, 2009.

[18] R. Beran, "Calibrating prediction regions," *Journal of the American Statistical Association*, vol. 85, no. 411, pp. 715–723, 1990.

[19] J. F. Lawless and M. Fredette, "Frequentist prediction intervals and predictive distributions," *Biometrika*, vol. 92, no. 3, pp. 529–542, 2005.

[20] A. V. Peterson Jr, "Bounds for a joint distribution function with fixed sub-distribution functions: Application to competing risks," *Proceedings of the National Academy of Sciences*, vol. 73, no. 1, pp. 11–13, 1976.

[21] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 2. The new algorithm," *IMA Journal of Applied Mathematics*, vol. 6, no. 3, pp. 222–231, 1970.

[22] R. Fletcher, "A new approach to variable metric algorithms," *The Computer Journal*, vol. 13, no. 3, pp. 317–322, 1970.

[23] D. Goldfarb, "A family of variable-metric methods derived by variational means," *Mathematics of Computation*, vol. 24, no. 109, pp. 23–26, 1970.

[24] D. F. Shanno, "Conditioning of quasi-newton methods for function minimization," *Mathematics of Computation*, vol. 24, no. 111, pp. 647–656, 1970.

[25] Y. Hong and W. Q. Meeker, "Field-failure predictions based on failure-time data with dynamic covariate information," *Technometrics*, vol. 55, no. 2, pp. 135–150, 2013.

[26] Z.-S. Ye, Y. Hong, and Y. Xie, "How do heterogeneities in operating environments affect field failure predictions and test planning?," *Annals of Applied Statistics*, vol. 7, pp. 2249–2271, 2013.

[27] Y. Hong, "On computing the distribution function for the poisson binomial distribution," *Computational Statistics & Data Analysis*, vol. 59, pp. 41–51, 2013.

Caleb King received MS in statistics (2011) from Virginia Tech. He is currently a PhD candidate in the Department of Statistics at Virginia Tech. His research interests include reliability data analysis, accelerated life testing, and small-sample distribution theory.

Yili Hong received BS in statistics (2004) from University of Science and Technology of China and MS and PhD degrees in statistics (2005, 2009) from Iowa State University. He is an Associate Professor in the Department of Statistics at Virginia Tech. His research interests include reliability data analysis, reliability test planning, and accelerated testing. He is an elected member of International Statistical Institute (ISI), and a member of American Statistical Association (ASA), Institute of Mathematical Statistics(IMS), and INFORMS. His work has been published in *Technometrics*, *IEEE Transactions on Reliability*, *Lifetime Data Analysis*, among others. He is currently an Associate Editor for *Technometrics*.

William Q. Meeker is a Professor of Statistics and Distinguished Professor of Liberal Arts and Sciences at Iowa State University. He is a Fellow of the American Statistical Association, and of the American Society for Quality. He is an elected member of the International Statistical Institute, and a past Editor of *Technometrics*. He is co-author of the books *Statistical Methods for Reliability Data* with Luis Escobar (1998), *Statistical Intervals: A Guide for Practitioners* with Gerald Hahn (1991), nine book chapters, and of numerous publications in the engineering and statistical literature. He has won the American Society for Quality (ASQ) Youden prize five times, the ASQ Wilcoxon Prize three times, and the American Statistical Association Best Practical Application Award. He is also an ASQ Shewhart Medalist. He has consulted extensively on problems in reliability data analysis, reliability test planning, accelerated testing, nondestructive evaluation, and statistical computing.