Review article

Institution of MECHANICAL ENGINEERS

Journal of RISK AND RELIABILITY

Proc IMechE Part O: J Risk and Reliability
1–19

DOI: 10.1177/1748006X251414181
journals.sagepub.com/home/pio
Sage

A mixed non-homogeneous Poisson process to model the quality of maintenance

Francisco Germán Badía[1], María Dolores Berrade[1], Hyunju Lee[2] and Iker Villegas[3]

Abstract
We present a model for the corrective and preventive maintenance of a system. The latter is based on a bivariate policy that replaces the system either at age T or after the M failure, whichever comes first. A repair follows each of the $M-1$ first failures, restoring the system operational state, but with a lower reliability than before failing. We present two scenarios with constant and time-dependent repair costs. The results reveal that systems with low initial reliability can greatly benefit from the bivariate policy. The advantage decreases for poor quality repairs. We also obtain conditions to obtain the optimum number M when T is given. This result is helpful to assess whether a system should be replaced sooner than originally planned.

Keywords
age replacement, mixture, non-homogeneous Poisson process, generalized Pólya process (GPP), optimum policy

Received: 8 April 2025; accepted: 17 December 2025

Introduction

Systems are usually repaired several times before being replaced. The extended time of use thus obtained must be weighed against the costs derived from failures and repairs to determine the optimum time for system renewal. Both, preventive and corrective maintenance are becoming increasingly important in the current world, where sustainability is driving the need to extend the life of all types of equipment.

The positive effect of maintenance on reliability is a general assumption in most models. It ranges from the minimal repair, that restores exactly the same reliability previous to failure, to system renewal and any other reliability level in between. However, the possibility of maintenance that leaves the system in a worse condition than before failing has received little attention in the literature. Cha et al.[1] justify this assumption by the negative effects of previous repairs, environmental and internal shocks, etc. Badía et al.[2] mention the lack of either adequate resources or properly trained maintainers as a reason for this poor maintenance.

The maintenance of roads and pipes shows this adverse effect. Potholes on highways are a serious safety risk to drivers, and they are constantly patched. It is also common to see the same point of the road being repaired repeatedly. This occurs because the patch is less reliable than the original pavement[3] or because the asphalt layer thickness is not optimal.[4] Khahro[5] highlights the impact of a low-cost pavement management. The increasing use of recycled materials in pavements[6] also leads to the assumption that these materials are less reliable when they are used in patches. Maintenance models have to assume the possibility that high-quality maintenance may not be carried out due to either limited budgets or studies that fail to detect the actual road deterioration.

Age replacement is a basic strategy for avoiding failures. Thus, the system is replaced when it fails or when it reaches a specified age, T, whichever comes first. However, some systems cannot be maintained

[1]Department of Statistics, School of Engineering and Architecture (EINA), University of Zaragoza, Spain
[2]Department of Statistics, Hankuk University of Foreign Studies, Seoul, South Korea
[3]A Former Student of Mathematics, Faculty of Science, University of Zaragoza, Spain

Corresponding author:
Francisco Germán Badía, Department of Statistics, School of Engineering and Architecture (EINA), University of Zaragoza, Zaragoza 50018, Spain.
Email: gbadia@unizar.es

periodically, but are better maintained at a random time, for example when a working cycle is completed,[7] when a software update is available[8] or after N failures.[9] The works of Mituzani et al.,[8] Zhao and Nakagawa,[10] and Badía et al.[2] consider the combination of random and periodic or age-based replacement.

The minimal repair assumption implies that the system recovers its functional state, but the failure rate remains undisturbed, equal to that of a system of the same age (as-bad-as-old). Potential industrial applications are found in systems consisting of a large number of components and a failed system is restored to the operating condition by replacing only the failed component. The non-homogeneous Poisson process (NHPP) models failures in systems undergoing minimal repairs.[11] It is useful to overcome the memoryless assumption when the deterioration process is not Markovian.[12] Overlooking heterogeneity between systems when analyzing failure data can lead to inaccurate estimations.[13] The generalized Pólya process (GPP) extends the NHPP by assuming that time varying environments affect the reliability of systems.[2,14] System functionality can be restored in different ways. For example, the component that replaced the failed one can be new or refurbished, or the maintenance team can have different levels of expertise. This variability in the quality of the repair is responsible for a heterogeneous population of systems and it must also be taken into account when designing maintenance policies; otherwise, they will not achieve the desired result. In particular, assuming that a minimal repair applies and, thus, a failure process following a NHPP, can overestimate the reliability of a system that is actually in a worse-than-old condition rather than as-bad-as-old. Unobserved heterogeneity between systems cannot be controlled and it is modeled by a frailty, Z. Z is a random variable which takes a particular value z for each system over its entire life, but changes between systems, explaining their differences. Therefore, assuming a mixture of non-homogeneous Poison processes (MNHPP) can be more realistic than using a single NHPP. In terms of managerial implications, a MNHPP can account for deficient maintenance and provide a dramatically different timescale for replacing a system. The relevance of considering heterogeneous populations in maintenance has recently been highlighted in Lee et al.[15] and Santos and Cavalcante.[16]

This paper presents two models that combine random and non-random maintenance of a system that undergoes minimal repairs after each of the first $M-1$ failures. The system is replaced after the first of the following two events occurs: the M failure or reaching age T. Thus, the system is no longer used when either its age or failure history induces high maintenance costs. This policy emulates the usual procedure of maintainers. Random replacement after M failures protects against frequent failures. This is particularly useful when failures happen in the early stages due to hidden defects that occur during the design stage or the manufacturing process or due to refurbished components.[17] Regarding road maintenance, it eliminates the potholes that appear in new roads due to a defective pavement or insufficient asphalt layer thickness.

The models in this paper differ from those in previous references in the following major assumptions:

1. We assume a MNHPP for the time to the M failure, which is an extension of the GPP repair process. The mixture is the result of differences in the quality of repairs. Thus, this new model accounts for random variability in repairing conditions, as well as time-varying reliability, which is more general than that in Badía et al.[2]
2. In model 2, we introduce time-dependent repair costs, $c_1(t)$, in contrast to the constant value, c_1, of model 1. This second assumption provides further insight into the comparison between deterministic and random preventive replacement.
3. From the point of view of real applications, we extend the idea of low quality maintenance in Santos and Cavalcante,[16] by considering successive repairs that restore the operational state of a system but with lower reliability than before the failure occurred (worse-than-old).[2] The consequence is that the occurrence of failures increases after each repair. This effect leads to the resurfacing of long sections of roads, for example. In addition, the bivariate maintenance policy is a useful strategy to determine the maximum usage time of a component subject to low-quality maintenance. The maintainer can decide when no further repairs are cost-effective and replace the component with a new one.

This paper is organized as follows: Section 2 presents the key concepts of the MNHPP, which models the repairs in models 1 and 2. Section 3 contains the model building that leads to the cost function under the two scenarios, constant repair cost and time-dependent repair cost. It also presents the analysis of the conditions under which an optimum policy exists. The corresponding proofs as well as some basics on stochastic ordering, are in the Appendices A1, A2 and A3. The classical univariate policies, $(T^\star, \infty)$ and $(\infty, M^\star)$, are presented in Section 4. The sensitivity analysis in Section 5 analyzes the range of application of models 1 and 2 when comparing the optimum bivariate policy, $(T^\star, M^\star)$, with the corresponding classical univariate, $(T^\star, \infty)$ and $(\infty, M^\star)$. The conclusions are summarized in Section 6.

MNHPP repair model

First, we introduce the notation that will be used throughout this paper:

$N(t)$	number of failures in $[0, t]$
$\lambda(t)$	failure rate of the time to the first failure
λ_t	stochastic intensity
Z	unobservable covariates (frailty)
$\lambda(t, z)$	failure rate conditional to $Z = z$
T_n	time to the n failure
M	maximum number of failures previous to replacement (decision variable)
T	time for preventive age replacement (decision variable)
τ	length of a renewal cycle
c_f	cost of renewal on the M failure
c_a	cost of age replacement at T
c_1	constant cost of repair in scenario 1
$c_1(t)$	time dependent cost of repair in scenario 2
$C(\tau)$	total cost of a cycle
$Q(T, M)$	cost rate (the long run expected cost per unit time)
DFR	decreasing failure rate
IFR	increasing failure rate

Let $\{N(t) : t \geqslant 0\}$ be a counting process with $N(t)$ the number of events in $[0, t]$. In addition $\mathfrak{H}_t$ represents its filtration, that is, the history of the process given by $N(t)$ and $T_1, \ldots T_{N(t^-)}$ with T_i the epoch time of the i th event in the counting process.

The stochastic intensity, λ_t, of a counting process $\{N(t) : t \geqslant 0\}$, is defined as

$$\lambda_t = \lim_{h \to 0} \frac{P(N(t, t + h) = 1 | \mathfrak{H}_t)}{h}$$

Observe that $h\lambda_t$ is the conditional probability of failure in an infinitesimal interval, h, given the previous history of failures of the system. Hence, the intensity rate is similar to the failure rate concept although λ_t accounts for the effect of the past failures in the current reliability of the system. In addition, assuming that repairing conditions cannot be perfectly controlled, but instead change randomly between systems, provides a more realistic model for failures, which are expected to occur more frequently in systems under low quality maintenance. The frailty, Z, accounts for such heterogeneity in the counting process of the failures. Z is assumed to be a non-negative random variable that leads to a mixture of counting processes. The work in Brown et al.[13] is concerned with the choice of the frailty.

In what follows we assume that the frailty, Z, has a multiplicative effect on the failure rate:

$$\lambda(t, z) = z\lambda(t)$$

$\lambda(t, z)$ is the failure rate when $Z = z$, whereas the failure rate of the time until the first failure, $\lambda(t)$, serves as baseline intensity. When $Z > 1$, the failure rate is higher than the baseline. On the contrary, $Z < 1$ leads to a reduction in the failure rate.

We consider the particular case of a MNHPP as in Cha and Finkelstein.[18] They provide an expression for the stochastic intensity (chapter 4, page 109) that, for the multiplicative model $\lambda(t, z) = z\lambda(t)$, turns out to be

$$\lambda_t = \lambda(t) \frac{E[Z^{N(t)+1} e^{-\Lambda(t)Z}]}{E[Z^{N(t)} e^{-\Lambda(t)Z}]} \tag{1}$$

with $\Lambda(t)$ being the cumulative baseline intensity:

$$\Lambda(t) = \int_0^t \lambda(x) dx \tag{2}$$

The stochastic intensity in equation (1) explains both the time aftermath, which is generally adverse for most systems, and random effects that can aggravate or alleviate the former. Note that for $Z = 1$ the expression in equation (1) corresponds to the intensity function of a non-homogeneous Poisson process which in turn describes the counting process of the failures for a system under minimal repair.

In what follows we will consider a system that is repaired after failing and keeps on working until the following failure. The mixed non-homogeneous repair process (MNHPP) is defined next.

Definition 1. *The counting process $\{N(t) : t \geqslant 0\}$ with $N(t)$ being the number of failures in $[0, t]$, is given by a MNHPP repair model with baseline intensity $\lambda(t)$ and frailty Z, if T_n $(n = 1, \ldots)$ corresponds to the n epoch time of the mixed non-homogeneous Poisson process with intensity λ_t in (1).*

The main property of the MNHPP repair process is given in the following proposition.

Proposition 1. *Let $\{N(t) : t \geqslant 0\}$ be a MNHPP repair model. Then its stochastic intensity, λ_t, increases with $N(t)$.*

The result holds since

$$\phi(n) = \frac{E[Z^{n+1} e^{-\Lambda(t)Z}]}{E[Z^n e^{-\Lambda(t)Z}]}, \ n = 0, 1, \ldots$$

is increasing in n.

Proof. Consider the ratio $\frac{\phi(n+1)}{\phi(n)}$ for $n = 0, 1, \ldots$:

$$\frac{\phi(n+1)}{\phi(n)} = \frac{E[Z^{n+2} e^{-\Lambda(t)Z}]}{E[Z^{n+1} e^{-\Lambda(t)Z}]} \frac{E[Z^n e^{-\Lambda(t)Z}]}{E[Z^{n+1} e^{-\Lambda(t)Z}]}$$

The Cauchy-Schwartz inequality states:

$$E^2[Z^{n+1} e^{-\Lambda(t)Z}] \leqslant E[Z^n e^{-\Lambda(t)Z}] E[Z^{n+2} e^{-\Lambda(t)Z}].$$

thus, $\frac{\phi(n+1)}{\phi(n)} \geqslant 1$, and the result follows.

Therefore, λ_t is also increasing with $N(t)$. The previous proposition implies that, although the system recovers its functionality after each repair, its condition is worse than before failing. Hence, the worse-than-old condition can be interpreted as the result of a mixture of minimal repairs in heterogeneous populations. This worse-than-old condition makes the difference between the MNHPP repair model and the minimal repair.

Furthermore, the MNHPP repair model is an extension of the GPP repair process considered in Lee and Cha.[19,20] In the particular case that the baseline failure rate in the NHPP be $\lambda(t) = \alpha r(t) e^{\int_0^t r(u)du}$ and the frailty, Z, a gamma random variable with scale parameter α and shapeparameter β/α, it follows that

$$E[Z^n e^{-Z\int_0^t \lambda(u)du}] = \frac{\Gamma(n+\beta/\alpha)}{\Gamma(\beta/\alpha)} \frac{\alpha^{\beta/\alpha}}{\left(\alpha + \int_0^t \lambda(u)du\right)^{n+\beta/\alpha}},$$

and then from equation (1) the stochastic intensity of the mixed NHPP process is

$$\lambda_t = \left(N(t) + \frac{\beta}{\alpha}\right) r(t) \tag{3}$$

which is also the stochastic intensity of the generalized Pólya process (GPP) with parameters $\left(r(t), 1, \frac{\beta}{\alpha}\right)$.

Model building

We consider a one component system undergoing a single type to failure that is minimally repaired after every failure with $\lambda(x)$ being the hazard rate of the time to the first failure. The higher $\lambda(x)$, the greater the risk of failure.

Definition 1 in Badía et al.,[2] states that the Generalized Pólya Process with parameters $(\lambda(t), \alpha = 0, \beta = 1)$ is a non-homogeneous Poisson process with rate equal to $\lambda(t)$. The corresponding probability of n failures in $(0, t)$ is a Poisson distribution with mean $\Lambda(t) = \int_0^t \lambda(x)dx$. The MNHPP in this paper and the corresponding repair process also extend the NHPP, which in turn emerges under the minimal repair. The probability of $N(t) = n$ failures $(n = 1, 2, \ldots)$ when $Z = z$ is

$$P(N(t) = n | Z = z) = \frac{(z\Lambda(t))^n}{n!} e^{-z\Lambda(t)}$$

That is, the probability of n failures in $(0, t)$ for a non-homogeneous Poisson process with rate equal to $z\lambda(t)$.

Then, the unconditional probability of $N(t) = n$ failures is

$$P(N(t) = n) = \int_0^\infty \frac{(z\Lambda(t))^n}{n!} e^{-z\Lambda(t)} \pi(z) dz$$

with $\pi(z)$ the density function of Z. Thus, the failure process, $\{N(t) : t \geq 0\}$, is given by a mixture of non-homogeneous Poisson processes.

Each time the system fails, a minimal repair restores the system back to function. Proposition 1 states that the random effect of the repair given by Z, leaves the system in a worse-than-old condition.

The following assumptions also apply to the maintenance procedure:

- The system is replaced on the M failure or at age T whichever comes first.
- Repair times are considered to be negligible.
- The cost of preventive replacement on the M failure and at age T are, respectively, c_f and c_a.

Although working times of most systems are finite, the study of optimum maintenance policies is simpler when an infinite time span is considered.[21] Since we do not assume that the system must fulfill a specific task within a given period, the objective function to be analyzed is the incurred cost per unit of time in an infinite interval. The key theorem of the renewal reward processes[22] guarantees that this function is equivalent to the next one:

$$Q(T, M) = \frac{E[C(\tau)]}{E[\tau]}$$

with τ and $C(\tau)$ being, respectively, the length and the cost of a renewal cycle. In addition T and M are decision variables. Therefore, given the costs of repair and replacement as well as the rest of parameters defining the failure process, the optimum values $T^\star$ and $M^\star$ minimizing $Q(T, M)$ have to be obtained.

Let S_n be the random time until the n-th failure. Its density function is given by:

$$f_{S_n}(x) = \int_0^\infty \frac{z^n \Lambda(x)^{n-1}}{(n-1)!} \lambda(x) e^{-z\Lambda(x)} \pi(z) dz \tag{4}$$

S_n and the process of failures, $\{N(t) : t \geq 0\}$, verify

$$P(S_M \leq t) = P(N(t) \geq M)$$

Hence the probability that the M failure occurs no later than t is equivalent to the probability that M failures occur at least before t.

Then, the length of a renewal cycle

$$\tau = \min(S_M, T)$$

and its expected value

$$E[\tau] = \int_0^T \bar{F}_{S_M}(t)dt = \sum_{k=0}^{M-1} \int_0^T P(N(t) = k)dt \tag{5}$$

with $\bar{F}_{S_M}(t) = P(S_M > t) = P(N(t) \leq M - 1)$.

In contrast with the renewal on the M failure, the renewal at age T is a planned maintenance. Therefore, the maintainer has time to prepare a less expensive replacement in advance than in the case of a random replacement, leading to consider $c_f > c_a$. However, since replacement on the M failure implies that the system has not reached age T, it may retain a residual value as a second-hand unit. Therefore, there may be cases where $c_f < c_a$.

We will analyze two scenarios:

- **Scenario 1:** The cost of repair is constant.
- **Scenario 2:** The costs of repairs are time dependent.

Scenario 1: Constant repair cost. Let c_1 be the unitary cost of repair.

The cost of a cycle is

$$C(\tau) = c_1 \min(M-1, N(T)) + c_a I_T + c_f I_M$$

$N(T)$ accounts for the number of failures occurring until T. I_T, I_M are the indicator variables of replacement at T or on the M failure, respectively.

The expected cost of a cycle is

$$\begin{aligned} E[C(\tau)] &= c_1 E[\min(M-1, N(T))] \\ &\quad + c_a P(I_T = 1) + c_f P(I_M = 1) \\ &= c_1 \sum_{k=0}^{M-2} k P(N(T) = k) \\ &\quad + c_1(M-1)P(N(T) \geqslant M-1) \\ &\quad + c_a P(S_M > T) + c_f P(S_M \leqslant T) \\ &= c_1 \left(M - 1 - \sum_{k=0}^{M-1} (M-1-k) P(N(T) = k) \right) \\ &\quad + c_a P(N(T) \leqslant M-1) + c_f P(N(T) \geqslant M) \\ &= c_1 \left(M - 1 - \sum_{k=0}^{M-1} (M-1-k) P(N(T) = k) \right) \\ &\quad + c_f + (c_a - c_f) \sum_{k=0}^{M-1} P(N(T) = k). \end{aligned} \tag{6}$$

The cost function in scenario 1 is

$$Q_1(T, M) = \frac{E[C(\tau)]}{E[\tau]}$$

with the mean cycle length, $E[\tau]$, and its corresponding expected cost, $E[C(\tau)]$, in equations (5) and (6), respectively.

Scenario 2: The costs of repairs are time dependent, with $c_1(t)$ the unitary cost.

Assume that n failures have happened in $(0, T)$, then the order statistics $S_1, \ldots, S_n | N(T) = n$, representing the corresponding times until they occur, are independent random variables with probability density function $\frac{\lambda(x)}{\Lambda(T)}$.

Moreover, consider a sample of size n from a random variable X, the density function of the ith order statistics, $X_{i:n}$, is:

For $0 \leqslant x \leqslant T$

$$\begin{aligned} f_{X_{i:n}}(x) &= \frac{n!}{(i-1)!(n-i)!} F_X(x)^{i-1} \bar{F}_X(x)^{n-i} f(x) \\ &= \frac{n!}{(i-1)!(n-i)!} \\ &\quad \times \left(\frac{\Lambda(x)}{\Lambda(T)} \right)^{i-1} \left(1 - \frac{\Lambda(x)}{\Lambda(T)} \right)^{n-i} \frac{\lambda(x)}{\Lambda(T)}. \end{aligned}$$

The renewal at T occurs if there is not more than $M-1$ failures in $[0, T]$. The corresponding probability is $\sum_{n=0}^{M-1} P(N(T) = n)$.

In addition, the times between consecutive failures, X_i, are independent random variables with density function $\frac{\lambda(x)}{\Lambda(T)}$.

The mean cost incurred due to the repair of n failures ($n \leqslant M-1$):

$$\begin{aligned} E\left[\sum_{i=1}^{n} c_1(S_i) | N(T) = n \right] &= \sum_{i=1}^{n} E[c_1(X_{i:n})] = \sum_{i=1}^{n} E[c_1(X_i)] \\ &= nE[c_1(X)] = n \int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx. \end{aligned}$$

The renewal after M failures happens in case that $S_M \leqslant T$. The corresponding probability is $\sum_{n=M}^{\infty} P(N(T) = n)$.

The mean cost derived from the repair of the first $M-1$ failures is as follows:

$$\begin{aligned} & E\left[\sum_{i=1}^{M-1} c(S_i) | N(T) = n \right] \\ &= \sum_{i=1}^{M-1} E[c(X_{i:n})] = \int_0^T c_1(x) n \sum_{j=0}^{M-2} \binom{n-1}{j} \times \\ & \left(\left(\frac{\Lambda(x)}{\Lambda(T)} \right)^j \left(1 - \frac{\Lambda(x)}{\Lambda(T)} \right)^{n-1-j} \frac{\lambda(x)}{\Lambda(T)} \right) dx. \end{aligned}$$

Then, the mean cost of a renewal cycle:

$$\begin{aligned} E[C(\tau)] &= \sum_{n=1}^{M-1} nP(N(T) = n) \int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx \\ &\quad + \sum_{n=M}^{\infty} P(N(T) = n) \int_0^T c_1(x) n \sum_{j=0}^{M-2} \binom{n-1}{j} \\ &\quad \times \left(\left(\frac{\Lambda(x)}{\Lambda(T)} \right)^j \left(1 - \frac{\Lambda(x)}{\Lambda(T)} \right)^{n-1-j} \frac{\lambda(x)}{\Lambda(T)} \right) dx \\ &\quad + c_f + (c_a - c_f) \sum_{k=0}^{M-1} P(N(T) = k). \end{aligned} \tag{7}$$

The cost function in scenario 2:

$$Q_2(T, M) = \frac{E[C(\tau)]}{E[\tau]}$$

with $E[\tau]$ and $E[C(\tau)]$ in equations (5) and (7), respectively.

Observe that in case of Z being a gamma random variable with scale parameter α and shape parameter β and $\lambda(t) = \alpha r(t)e^{\alpha \int_0^t r(u)du}$, then the cost functions in scenarios 1 and 2 extend that in Lee and Cha[20] assuming $c_a = c_f = c_r$ and $c_1 = c_{GPP}$ in the case of scenario 1 and $c_1(t) = c_{GPP}$ for scenario 2.

We aim at obtaining the optimum policy in both scenarios, that is, $(T_1^\star, M_1^\star)$ minimizing $Q_1(T, M)$ and the corresponding $(T_2^\star, M_2^\star)$ for $Q_2(T, M)$.

Optimum number of repairs, M

This section concerns the existence of an optimum $M^\star$ minimizing $Q(T, M)$, for a given T, in scenarios 1 and 2. The first result provides the algorithm to obtain $M^\star$, simplifying programming tasks. The second serves to explore conditions under which there exists $M^\star$.

The notation below is used in what follows

$$C(T, M) = E[C(\tau)].$$

$$L(T, M) = C(T, M)\frac{\int_0^T P(N(t) = M)dt}{C(T, M + 1) - C(T, M)} - \sum_{j=0}^{M-1} \int_0^T P(N(t) = j)dt.$$

The following two results show the applicability of this model. First, Proposition 2 presents the strategy for computing the optimum M, if it exists, when T is given. It applies in both scenarios.

Proposition 2. Let $M^* = arg_M \min Q(T, M)$ for a given $T > 0$. If the function given next

$$\frac{\int_0^T P(N(t) = M)dt}{C(T, M + 1) - C(T, M)} \tag{8}$$

is decreasing in M, and it verifies the following condition

$$\lim_{M\to\infty} \frac{\int_0^T P(N(t) = M)dt}{C(T, M + 1) - C(T, M)} = 0 \tag{9}$$

then

$$M^\star = \inf\{M > 1 : L(T, M) < 0\}$$

if $L(T, 1) > 0$. *If* $L(T, 1) \leqslant 0$, *then* $M^* = \infty$.

The proof is in the Appendix A2.

The next theorem states sufficient conditions for the existence of an optimum M in each scenario when the age replacement time is given. This information keeps the maintenance team aware of systems that need to be replaced earlier than planned.

Theorem 1. *For a given age replacement time T, the following results hold:*

a) *Consider* $Q_1(T, M)$ *in scenario 1 with* $c_1 + c_a - c_f > 0$. *If either of the following conditions applies*

a_1) Z is log concave and $\lambda(x)$ is increasing in x

a_2) Z is DFR, $\lambda(x)$ increasing in x and $\frac{\pi(az)}{\pi(z)}$ is increasing in z for $0 \leqslant a \leqslant 1$

then, there exists $M_1^\star = arg_M \min Q_1(T, M)$.

b) *Consider* $Q_2(T, M)$ *in scenario 2 with* $c_a = c_f$. *If* Z *is IFR,* $c_1(x)$ *and* $\lambda(x)$ *are increasing in x and* $\frac{\pi(az)}{\pi(z)}$ *is increasing in z for* $0 \leqslant a \leqslant 1$

then, there exists $M_2^\star = arg_M \min Q_2(T, M)$.

The proof is in the Appendix A3.

Observe that if Z is a gamma random variable with scale parameter α and shapeparameter $\frac{\beta}{\alpha}$, then Z is log concave for $\beta \geqslant \alpha$ and DFR otherwise. Therefore, if $\lambda(x)$ is increasing in x and $c_1 + c_a - c_f > 0$ the conditions in a_1) apply for $\beta \geqslant \alpha$.

In addition, its density function verifies that $\frac{\pi(az)}{\pi(z)}$ is increasing in z for $0 \leqslant a \leqslant 1$ and, hence, the conditions in a_1) hold. Thus, conditions in Theorem 1 for scenario 1 extend the results in Lee and Cha[20] (see Proposition 3).

Theorem 1 has interesting applications for maintainers when the age for replacement of a system, T, is given. For example, this is useful when T represents an amortization period. The corresponding bivariate policy takes into account the resulting reliability of the system after successive repairs. In the case of low quality repairs, the optimum policy may be an earlier replacement than indicated by the amortization period. Regarding the sufficient condition $c_1 + c_a - c_f > 0$, it seems to be a reasonable assumption when inspection is an expensive procedure. In addition, it trivially holds if $c_a \geqslant c_f$. As mentioned previously, cases in which the system retains a residual value as a second-hand unit can imply that $c_a > c_f$.

Comparison with univariate policies

In this section we present the classical univariate maintenance models, age replacement (T, ∞) and replacement after a given number of failures (∞, M) assuming a MNHPP repair model. A comparison of the optimum cost derived from the three models gives light about the conditions under which the bivariate maintenance outperforms the other two.

Next we present the corresponding expressions of the cost function in scenarios 1 and 2.

Age replacement at T. M = ∞

The length of a cycle is $\tau = T$ (non-random). The mean number of repairs is that of the mixture of non-homogeneous Poisson processes. In addition to the replacement cost, the cost derived from repairs until T have to be considered. It follows that

$$E[\tau] = T.$$

The cost of a cycle in scenarios 1 and 2 are respectively:

$$E[C(\tau)|s_1] = c_1 E[N(T)] + c_a = c_1 \Lambda(T) E[Z] + c_a.$$

$$E[C(\tau)|s_2] = E[N(T)] \int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx + c_a$$

$$= \Lambda(T) E[Z] \int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx + c_a.$$

with $Q_{T,s_1}(T)$ and $Q_{T,s_2}(T)$ being the corresponding cost functions in the two scenarios.

The optimum values minimizing $Q_{T,s_1}(T)$ and $Q_{T,s_2}(T)$ are denoted, respectively, by $T^\star_{s_1}$ and $T^\star_{s_2}$.

Replacement after M failures. $T = \infty$

The cycle verifies $\tau = S_M$ and therefore it presents a random length, but the incurred cost is deterministic. It follows that

$$E[\tau] = E[S_M] = \int_0^\infty x f_{S_M}(x) dx$$

with $f_{S_M}(x)$ in (4).

The length of a cycle can be alternatively expressed as follows:

$$E[\tau] = E[S_M] = \int_0^\infty P(N(x) \leqslant M-1) dx$$

$$= \sum_{k=0}^{M-1} \int_0^\infty P(N(x) = k) dx.$$

The corresponding cost in scenario 1:

$$E[C(\tau)|s_1] = (M-1)c_1 + c_f.$$

and the cost function

$$Q_{M,s_1}(M) = \frac{c_1(M-1) + c_f}{E[S_M]}.$$

Regarding scenario 2 and $E[C(\tau)]$ in equation (7), the expected cost of repairing a failure in $(0, T)$ is

$$\int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx.$$

Observe that $H(x) = \frac{\Lambda(x)}{\Lambda(T)}$ verifies that $dH(x) = \frac{\lambda(x)}{\Lambda(T)} dx$. Considering $u = H(x)$, it follows that

$$\int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx = \int_0^1 c_1\left(H^{-1}(u)\right) du.$$

Replacement on the M failure when there is at least one repair ($M > 1$) makes no sense if the expected cost of repair is infinite. The assumption below is required for the expected cost derived from repairs in a cycle to be finite in scenario 2:

$$\lim_{T \to \infty} \int_0^T c_1(x) \frac{\lambda(x)}{\Lambda(T)} dx$$

$$= \lim_{T \to \infty} \int_0^1 c_1\left(H^{-1}(u)\right) du = \lim_{T \to \infty} c_1(T) < \infty.$$

and, thus, the expected cost of a cycle is

$$E[C(\tau)|s_2] = (M-1) \int_0^1 c_1\left(H^{-1}(u)\right) du + c_f.$$

The cost function in scenario 2:

$$Q_{M,s_2}(M) = \frac{(M-1) \int_0^1 c_1\left(H^{-1}(u)\right) du + c_f}{E[S_M]}.$$

Now $M^\star_{s_1}$ and $M^\star_{s_2}$ denote the optimum number of failures previous to replacement minimizing $Q_{M,s_1}(M)$ and $Q_{M,s_2}(M)$, respectively.

Aiming at giving additional insight about the advantages of a bivariate policy, $Q_1\left(T^\star, M^\star\right)$ and $Q_2\left(T^\star, M^\star\right)$ have to be compared with the univariate policies in both scenarios.

Numerical example

In this section we carry out a sensitivity analysis for both scenarios. We obtain and compare the corresponding optimum policies for the univariate and bivariate maintenance. Thus, we give light on how the decision variables depend on the parameters of the model, as well as about the conditions under which the cost of the univariate maintenance is no longer larger than that of the bivariate policy. It is important to note that the use of a bivariate strategy for replacement adds extra complexity for maintainers. Therefore, these results are useful to indicate when a simpler maintenance based on just one of the two can be applied.

Table 1. Optimum Policies in scenario 1 ($c_1 = 50$).

Case	c_f	b	s	$T^\star$	$M^\star$	$E[\tau]$	$Q_1(T^\star, M^\star)$	$T_1^\star$	$Q_{T,s_1}(T_1^\star)$	$\Delta C_{T,s_1}\%$	$M_1^\star$	$Q_{M,s_1}(M_1^\star)$	$\Delta C_{M,s_1}\%$
1	120	1.5	0.5	∞	1	11.812	10.159	12.460	12.425	18.241	1	10.159	0
2			1.0	∞	1	9.305	12.896	10.355	16.146	20.129	1	12.896	0
3			1.5	∞	1	7.966	15.064	9.174	19.111	21.177	1	15.064	0
4			3.0	∞	2	8.602	19.762	7.287	26.156	24.445	2	19.762	0
5		3.0	0.5	∞	1	8.507	14.105	10.355	16.146	12.642	1	14.105	0
6			1.0	17.412	2	9.165	18.468	8.368	21.686	14.837	2	18.474	0.032
7			1.5	∞	2	7.773	21.865	7.287	26.156	16.405	2	21.865	0
8			3.0	∞	2	5.651	30.081	5.621	36.913	18.506	2	30.081	0
9	130	1.5	0.5	24.416	1	11.761	11.000	12.460	12.425	11.469	1	11.005	0.047
10			1.0	∞	1	9.305	13.971	10.355	16.146	13.473	1	13.971	0
11			1.5	19.669	2	10.974	16.140	9.174	19.111	15.544	2	16.165	0.155
12			3.0	∞	2	8.598	20.925	7.287	26.156	20.003	2	20.925	0
13		3.0	0.5	13.985	2	11.075	14.894	10.355	16.1456	7.754	2	15.133	1.581
14			1.0	14.012	2	8.980	19.504	8.368	21.686	10.061	2	19.561	0.29
15			1.5	15.406	2	7.732	23.142	7.287	26.156	11.522	2	23.152	0.041
16			3.0	∞	2	5.651	31.851	5.621	36.913	13.712	2	31.851	0
17	140	1.5	0.5	15.747	2	13.587	11.665	12.460	12.425	6.116	1	11.852	1.575
18			1.0	15.659	2	11.773	14.698	10.355	16.146	8.966	2	14.918	1.474
19			1.5	16.749	2	10.714	16.977	9.174	19.111	11.166	2	17.064	0.506
20			3.0	22.543	2	8.582	22.083	7.287	26.156	15.572	2	22.087	0.019
21		3.0	0.5	12.803	2	10.659	15.487	10.355	16.146	4.084	2	15.974	3.051
22			1.0	12.086	2	8.684	20.457	8.368	21.686	5.667	2	20.648	0.924
23			1.5	12.440	2	7.576	24.374	7.287	26.156	6.813	2	24.438	0.261
24			3.0	∞	2	5.650	33.620	5.621	36.913	8.920	2	33.620	0
25	150	1.5	0.5	14.842	2	12.393	12.031	12.460	12.425	3.173	1	12.698	5.257
26			1.0	14.239	2	11.331	15.292	10.355	16.146	5.288	2	15.703	2.618
27			1.5	14.711	2	10.363	17.760	9.174	19.111	7.070	2	17.962	1.123
28			3.0	19.378	2	8.039	23.234	7.287	26.156	11.173	2	23.250	0.070
29		3.0	0.5	11.122	3	10.633	15.966	10.355	16.146	1.116	2	16.815	5.049
30			1.0	9.567	3	8.790	21.266	8.368	21.686	1.934	2	21.7342	2.153
31			1.5	8.862	3	7.820	25.471	7.287	26.156	2.620	2	25.7240	0.985
32			3.0	8.511	3	6.416	35.315	5.621	36.913	4.329	2	35.3898	0.212
33	175	1.5	0.5	12.507	6	12.491	12.424	12.460	12.425	0.010	2	14.4294	13.901
34			1.0	10.575	5	10.496	16.122	10.355	16.146	0.147	2	17.6661	8.739
35			1.5	9.776	4	9.502	19.032	9.174	19.111	0.414	2	20.2068	5.815
36			3.0	9.603	3	8.210	25.720	7.287	26.156	1.665	2	26.1559	1.665
37		3.0	0.5	10.351	∞	10.351	16.146	10.355	16.146	0	2	18.9165	14.646
38			1.0	8.368	∞	8.368	21.686	8.368	21.686	0	2	24.4510	11.310
39			1.5	7.287	∞	7.287	26.156	7.287	26.156	0	2	28.9395	9.618
40			3.0	5.654	8	5.643	36.907	5.621	36.913	0.016	3	39.1216	5.662

The baseline hazard rate is given as follows

$$\lambda(t|s) = 0.01s(t+1)e^{\int_0^t 0.01(u+1)du}$$

with $s > 0$. If $s_1 < s_2$ then $\lambda(t|s_1) < \lambda(t|s_2)$. Therefore, the parameter s models the pace of the first failure. The higher s, the more prone the system is to an early failure. This can be the result of different causes such as hidden defects, poor installation of the system, refurbished units used as spare parts, etc.

The mixing variable Z follows a gamma distribution with density function:

$$\pi(z) = \frac{a^b z^{b-1} e^{-az}}{\Gamma(b)}, \quad z > 0. \tag{10}$$

with $a > 0$, $b > 0$ and $\Gamma(b)$ the Euler function:

$$\Gamma(b) = \int_0^\infty e^{-t} t^{b-1} dt.$$

In what follows we will assume $a = 1$ and $c_a = 100$. Tables 1 and 2 are obtained for $c_1 = 50$ and $c_1(t) = 50 + 30(1 - e^{-0.1t})$ corresponding to scenario 1 and 2, respectively. Both contain the optimum bivariate policy $(T^\star, M^\star)$, the optimum cost $Q_i(T^\star, M^\star)$ as well as the univariate policies $T_i^\star$ and $M_i^\star$ along with the corresponding optimum costs, $Q_{T,s_i}(T_i^\star)$ and $Q_{M,s_i}(M_i^\star)$, for $i = 1$ (scenario 1) in Table 1 and $i = 2$ (scenario 2) in Table 2. Table 1 also contains the following information for comparison purposes with the univariate policies:

Table 2. Optimum Policies in scenario 2 ($c_a = 100$).

Case	c_f	b	s	T	M	$E[\tau]$	$Q_2(T, M)$	T_2	$Q_{T,s_2}(T_2^\star)$	$\Delta C_{T,s_2}\%$	$M_2^\star$	$Q_{M,s_2}(M_2)$	$\Delta C_{M,s_2}\%$
1	100	1.5	0.5	∞	1	11.813	8.466	11.296	13.738	38.379	1	8.466	0
2			1.0	∞	1	9.305	10.747	9.306	17.835	39.746	1	10.747	0
3			1.5	∞	1	7.966	12.553	8.212	21.058	40.388	1	12.553	0
4			3.0	∞	1	5.943	16.826	6.498	28.607	41.182	1	16.826	0
5		3.0	0.5	∞	1	8.508	11.754	9.306	17.835	34.097	1	11.754	0
6			1.0	∞	1	6.335	15.785	7.474	23.833	33.769	1	15.785	0
7			1.5	∞	1	5.240	19.084	6.498	28.607	33.288	1	19.084	0
8			3	∞	1	3.680	27.176	5.018	39.944	31.967	1	27.176	0
9	125	1.5	0.5	27.680	1	11.800	10.581	11.296	13.738	22.979	1	10.582	0.008
10			1.0	∞	1	9.305	13.433	9.306	17.835	24.682	1	13.433	0
11			1.5	∞	1	7.966	15.691	8.212	21.058	25.486	1	15.692	0
12			3.0	∞	1	5.943	21.033	6.498	28.607	26.478	1	21.033	0
13		3.0	0.5	20.528	1	8.499	14.691	9.306	17.835	17.629	1	14.693	0.01
14			1.0	∞	1	6.335	19.731	7.474	23.833	17.211	1	19.731	0
15			1.5	∞	1	5.240	23.855	6.498	28.607	16.61	1	23.855	0
16			3	17.013	2	5.649	32.367	5.018	39.945	18.97	1	33.969	4.717
17	150	1.5	0.5	18.376	1	11.364	12.594	11.296	13.738	8.324	1	12.698	0.818
18			1.0	20.469	1	9.217	16.102	9.306	17.835	9.72	1	16.120	0.112
19			1.5	23.705	1	7.952	18.827	8.212	21.058	10.596	1	18.830	0.015
20			3.0	17.013	2	8.452	24.608	6.498	28.607	13.978	1	25.239	2.499
21		3.0	0.5	11.248	2	9.915	17.043	9.306	17.835	4.442	1	17.631	3.335
22			1.0	10.098	2	8.106	22.533	7.474	23.833	5.452	1	23.677	4.83
23			1.5	9.783	2	7.136	26.849	6.498	28.607	6.146	1	28.627	6.21
24			3.0	11.702	2	5.591	36.893	5.018	39.945	7.64	2	40.698	9.351
25	175	1.5	0.5	12.651	2	11.753	13.581	11.296	13.738	1.143	1	14.815	8.328
26			1.0	11.339	2	9.981	17.502	9.306	17.835	1.867	1	18.807	6.935
27			1.5	10.884	2	9.072	20.521	8.212	21.058	2.551	1	21.968	6.587
28			3.0	11.790	2	7.854	27.290	6.498	28.607	4.603	1	29.446	7.321
29		3.0	0.5	9.438	4	9.379	17.814	9.306	17.835	0.118	1	20.570	13.395
30			1.0	7.678	4	7.582	23.789	7.474	23.833	0.184	1	27.623	13.881
31			1.5	6.753	4	6.629	28.539	6.498	28.607	0.238	2	32.798	12.986
32			3.0	5.389	4	5.201	39.792	5.018	39.9445	0.382	2	45.122	11.812
33	200	1.5	0.5	11.296	∞	11.296	13.738	11.296	13.7379	0	1	16.931	18.859
34			1.0	9.326	7	9.322	17.834	9.306	17.835	0.007	1	21.493	17.024
35			1.5	8.275	6	8.258	21.053	8.212	21.058	0.025	1	25.106	16.145
36			3.0	6.723	5	6.649	28.577	6.498	28.607	0.104	2	32.550	12.205
37		3.0	0.5	9.306	∞	9.306	17.835	9.306	17.835	0	1	23.508	60.412
38			1.0	7.474	∞	7.474	23.833	7.474	23.833	0	2	30.428	21.675
39			1.5	6.498	∞	6.498	28.607	6.498	28.607	0	2	36.013	20.566
40			3.0	5.018	∞	5.018	39.945	5.018	39.945	0	2	49.546	19.378

$$\Delta C_{T,s_1}\% = \frac{Q_{T,s_1}(T_1^\star) - Q_1(T^\star, M^\star)}{Q_{T,s_1}(T_1^\star)},$$
$$\Delta C_{M,s_1}\% = \frac{Q_{M,s_1}(M_1^\star) - Q_1(T^\star, M^\star)}{Q_{M,s_1}(M_1^\star)}.$$

to compare with age replacement and replacement on the M failure, respectively. Table 2 presents a similar study in scenario 2. The larger any of the two quantities above, the greater the advantage of the bivariate over its univariate competitors.

As c_f/c_a increases, so does $M^\star$, but $T^\star$ decreases. The higher cost of replacement on failure impels age replacement by reducing T and increasing M. Therefore, the relevance of the age replacement increases. Under this condition, the finite value of $T^\star$ prevents the system from exceeding M. Very large values of the ratio c_f/c_a can result in $M^\star = \infty$ (cases 37–39 in Table 1 and 37–40 in Table 2). On the contrary, as the ratio c_f/c_a decreases, replacement based on the accumulated number of failures becomes more profitable than age replacement at T, and therefore $T^\star = \infty$ is observed more frequently. In fact, the cases considered where $c_f \simeq c_a$ lead to $T^\star = \infty$.

The higher s, so is the initial failure rate of the system, leading to a decrease in expected length of a renewal cycle. The stochastic intensity of the failure process increases with b and also leads to a reduction in $E[\tau]$. Therefore, the starting condition of a recycled component is critical when considering its use as a second-hand system. Poor repairs can also interfere with the purpose of extending the life of a system. Figure 1 illustrates this result, which must be considered when either recycled units are used as spares or the quality level of repairs drops. The comparison of the bivariate policy with the age replacement in scenarios 1 and 2 reveals that the cost reduction induced by the former decreases with b. This means that the additional replacement on the M failure is less advantageous the

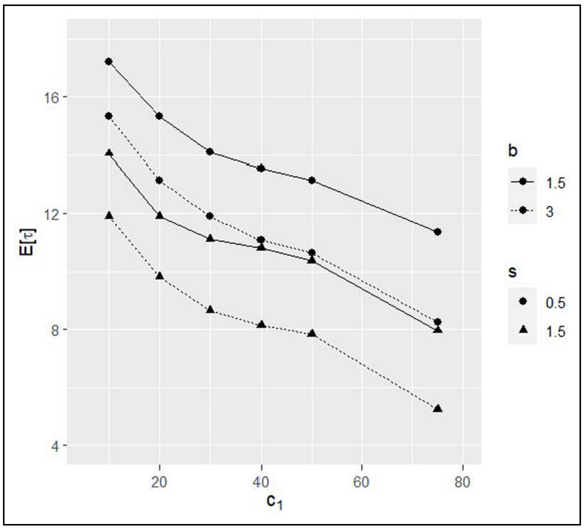

Figure 1. Mean length of a renewal cycle under different values of b, s, and c_1.

lower the quality of repairs is. This result is reversed in both scenarios when s increases. Therefore, the bivariate policy is clearly superior to the age replacement in systems with low reliability when they start to work.

Notwithstanding that the bivariate policy outperforms the univariate maintenances in most cases, the sensitivity analysis also gives some insight into the conditions leading to either, $T^\star$ or $M^\star$ being infinite. If so, the optimum cost of the univariate policy is equal to that of the bivariate policy. Table 3 provides further information on the interaction effects of the three parameters, c_f, b, and s in determining the optimum policy in scenario 2. If b and s are fixed, $M^\star < \infty$ for $c_f = 190$ and $M^\star = \infty$ for $c_f = 200$ (cases 5, 9, 13, 14, and 15). When b is fixed, increasing s leads to a decrease in $M^\star$, with some cases where $M^\star = \infty$ drops to $M^\star < \infty$. This behavioris reversed when s is fixed and b increases. Figure 2 shows these results. Thus, the additional protection of the bivariate policy is more advantageous the lower the initial reliability of the system is. The bivariate policy tends to match an age replacement for large values of the parameter b, which is the shape parameter of the mixing distribution Z. According to equation (3), the larger b, the worse the repair. In summary, a good maintenance is key for keeping systems operational. However, in case of a very low quality of repairs, the results do not show a clear gain in a more complex maintenance. An earlier age replacement might be sufficient.

Conclusions

Replacing systems rather than repairing them when they fail or deteriorate, has been a common practice since the last decade of the last century. Nonetheless, this practice is no longer acceptable for sustainability reasons that lead to extend the use of components and systems by appropriate maintenance. This paper presents a bivariate maintenance policy based on both the age, T, and the number of failures, M. The combination of a deterministic and a random replacement generally gives additional protection to the system, since waiting for the age replacement at T, may not be worthwhile in a system with a high frequency of failures. Therefore, a preventive maintenance after the M failure, which

Table 3. Optimum Policies in scenario 2 ($c_a = 100$).

	$c_f = 90$				$c_f = 200$					
Case	b	s	$T^\star$	$M^\star$	$E[\tau]$	$Q_2(T^\star, M^\star)$	$T^\star$	$M^\star$	$E[\tau]$	$Q_2(T^\star, M^\star)$
1	1.25	0.5	12.261	3	12.062	12.864	11.863	6	11.857	12.887
2		1.0	10.519	3	10.202	16.526	9.937	5	9.903	16.595
3		1.5	9.633	3	9.216	19.364	8.874	5	8.826	19.494
4		3.0	8.510	3	7.797	25.861	7.517	4	7.308	26.203
5	1.50	0.5	11.443	4	11.387	13.729	11.296	∞	11.296	13.738
6		1.0	9.570	4	9.474	17.802	9.326	7	9.322	17.834
7		1.5	8.570	4	8.441	20.991	8.275	6	8.258	21.053
8		3.0	7.726	3	7.097	28.392	6.723	5	6.649	28.577
9	1.75	0.5	10.893	5	10.876	14.520	10.843	∞	10.843	14.522
10		1.0	8.983	5	8.952	18.964	8.886	9	8.885	18.976
11		1.5	7.954	5	7.910	22.462	7.823	8	7.820	22.488
12		3.0	6.609	4	6.416	30.643	6.194	7	6.182	30.733
13	2.00	0.5	10.471	6	10.466	15.255	10.455	∞	10.455	15.256
14		1.0	8.560	6	8.550	20.042	8.524	∞	8.524	20.046
15		1.5	7.529	6	7.514	23.823	7.474	∞	7.474	23.833
16		3.0	6.043	5	5.974	32.711	5.858	10	5.857	32.750
17	3.00	0.5	9.307	∞	9.307	17.836	9.306	∞	9.306	17.835
18		1.0	7.474	∞	7.474	23.833	7.474	∞	7.474	23.833
19		1.5	6.498	∞	6.498	28.607	6.498	∞	6.498	28.607
20		3.0	5.018	∞	5.018	39.945	5.018	∞	5.018	39.945

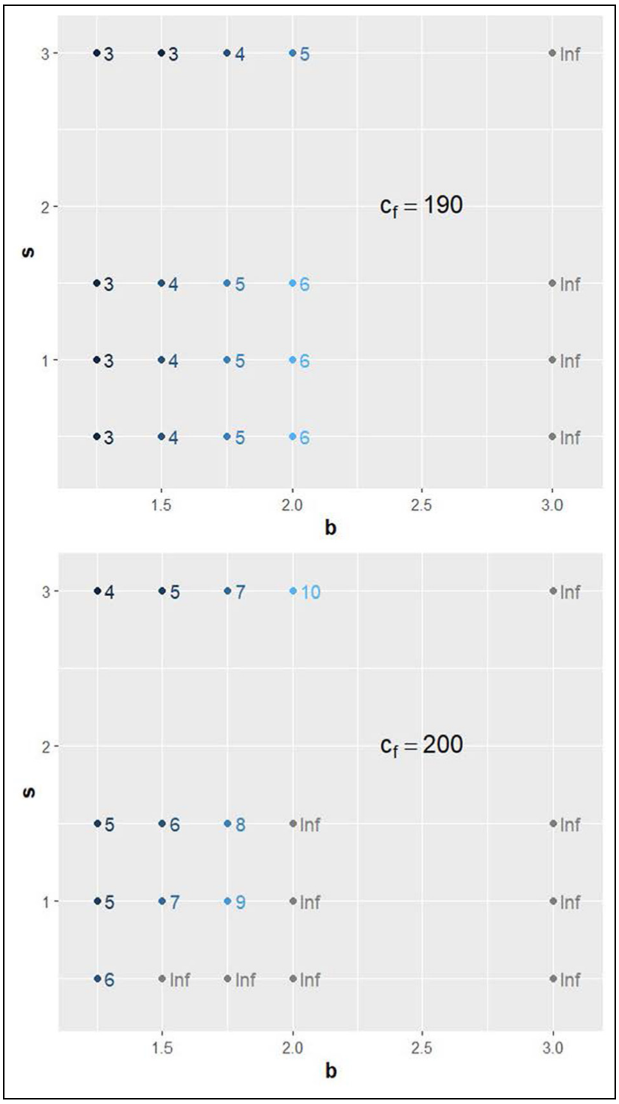

Figure 2. $M^{\star}$ in scenario 2.

occurs before the planned age replacement at T, can be profitable from an economic point of view. However, poor quality maintenance can reduce the useful life of a system. There is unobserved heterogeneity between systems that seem to be identical due to differences in the quality of their maintenance. This variability is modeled by a frailty, Z, which leads to a MNHPP in the failure process, which, in turn, is derived from a variety of minimal repairs carried out on failure. The value of Z is specific to each system and remains constant throughout its life, but varies between systems depending on the quality of maintenance. This paper focuses on designing a maintenance policy that is valid for a heterogeneous population of systems. We analyze the consequences of low quality maintenance in the optimum policies under two scenarios with constant or time-dependent repair costs, respectively. The higher these costs, the earlier the system replacement implying a serious inconvenience to extend the time of use. We also provide sufficient conditions guaranteeing the existence of a finite M in each of the two scenarios when T is not a variable decision but a parameter in the model. This result, the maximum number of cost-effective repairs to be carried out, can serve as an indicator that an earlier replacement of the system, that is before the initial horizon at T, iseconomically advantageous. Apart from the sensitivity analysis, the numerical study also addresses the comparison of the bivariate policy with the univariate ones, based only on the age $\left(M^{\star} = \infty\right)$ or the number of cumulated failures $\left(T^{\star} = \infty\right)$. The initial reliability of a system when it starts to work is critical to determine the upcoming maintenance. The lower the initial reliability, the sooner the system will be replaced. Thus, the condition of a component that is to be used as a second-hand system is relevant to determine how it should be maintained. In fact, recycled units with low reliability levels can greatly benefit from the bivariate policy. The analysis reveals that some conditions can lead maintainers to use univariate maintenance policies, which are simpler to apply. Thus, $T^{\star} = \infty$ is more likely to occur when the ratio c_f/c_a is low, whereas $M^{\star} = \infty$ tends to happen either when the former ratio is high or the mean quality of repairs decreases.

Acknowledgments

The authors thank the anonymous reviewers whose helpful comments significantly improved the final version of this paper.

ORCID iD

Francisco Germán Badía https://orcid.org/0000-0002-6651-3306

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The work of F. G. Badía and M. D. Berrade is supported by the Spanish Ministry of Science and Innovation and the Spanish National Agency for Research under Projects PID2021-123737NB-I00 and PID2024-155364NB-I00, respectively. The work of H. Lee was supported by Hankuk University of Foreign Studies Research Fund of 2024 and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2023-00240817).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Cha JH, Finkelstein MS and Levitin G. On the delayed worse-than-minimal repair model and its application to preventive replacement. *IMA J Manag Math* 2023; 34(1): 101–122.

2. Badía FG, Berrade MD, Cha JH, et al. Optimal replacement policy under a general failure and repair model: minimal versus worse than old repair. *Rel Eng Syst Saf* 2018; 180: 362–372.
3. Wang T, Dra YASS, Cai XP, et al. Advanced cold patching materials (CPMs) for asphalt pavement pothole rehabilitation: state of the art. *J Clean Prod* 2022; 366: Art. no. 133001.
4. Zouch M, Yeung TG and Castanier B. Optimizing road milling and resurfacing actions. *Proc Inst Mech Eng O J Risk Reliab* 2012; 226(O2): 156–168.
5. Khahro SH. Defects in flexible pavements: a relationship assessment of the defects of a low-cost pavement management system. *Sustainability* 2022; 14(24): 16475.
6. Huang Y, Bird R and Heidrich O. Development of a life cycle assessment tool for construction and maintenance of asphalt pavements. *J Clean Prod* 2009; 17(2): 283–296.
7. Nakagawa T. *Maintenance theory of reliability*. Springer, 2005.
8. Mituzani S, Zhao X and Nakagawa T. Age and periodic replacement policies with two failure modes in general replacement models. *Rel Eng Syst Saf* 2021; 214: Art. no. 107754.
9. Zong S, Chai G, Zhang ZG, et al. Optimal replacement policy for a deteriorating system with increasing repair times. *Appl Math Model* 2013; 37: 9768–9775.
10. Zhao X and Nakagawa T. Optimization problems of replacement first or last in reliability theory. *Eur J Oper Res* 2012; 223: 141–149.
11. Høyland A and Rausand M. *System reliability theory. Models and statistical methods*. Wiley, 1994.
12. Safaei F and Taghipour S. An availability-constrained integrated maintenance–monitoring model for a system with failures following an NHPP. *IEEE Trans Rel* 2024; 73(2): 937–951.
13. Brown B, Liu B, McIntyre S, et al. Reliability evaluation of repairable systems considering component heterogeneity using frailty model. *Proc Inst Mech Eng O J Risk Reliab* 2023; 237(4): 654–670.
14. Liu P and Wang GJ. Generalized polya-process-based reliability analysis for systems working under dynamic environment. *IEEE Trans Rel* 2024; 74(1): 2146–2156.
15. Lee KL, Lan LH, Chien YH, et al. Probabilistic and cost analyses of a renewable warranty with an inspection policy for a discrete operating item from a heterogeneous population. *Appl Math Model* 2021; 100: 138–151.
16. Santos ACD and Cavalcante CAV. A study on the economic and environmental viability of second-hand items in maintenance policies. *Rel Eng Syst Saf* 2022; 217: Art. no. 108133.
17. Berrade MD, Calvo E and Badía FG. Maintenance of systems with critical components. Prevention of early failures and wear-out. *Comput Ind Eng* 2023; 181: Art. no. 109291.
18. Cha JH and Finkelstein MS. *Point processes for reliability analysis*. Springer, 2018.
19. Lee H and Cha JH. New stochastic models for preventive maintenance optimization. *Eur J Oper Res* 2016; 255: 80–90.
20. Lee H and Cha JH. A bivariate optimal replacement policy for a system subject to a generalized failure and repair process. *Appl Stoch Model Bus* 2019; 35: 637–650.
21. Nakagawa T and Finkelstein S. A summary of maintenance policies for a finite interval. *Rel Eng Syst Saf* 2009; 94: 89–96.
22. Ross SM. *Stochastic processes*. 2nd ed. Wiley, 2008.
23. Shanthikumar JG and Yao DD. Bivariate characterization of some stochastic order relations. *Adv Appl Probab* 1991; 23: 642–659.
24. Cuadras CM. On the covariance between functions. *J Multivar Anal* 2002; 81: 19–27.
25. Joe H. *Multivariate models and dependence concepts*. Chapman & Hall, 1997.
26. Grandell J. *Mixed Poisson processes*. Chapman & Hall, 1997.
27. Shaked M and Shanthikumar JG. *Stochastic orders*. Springer, 2007.

Appendix A1. Basics

Consider a random variable X

- F_X, $\bar{F}_X$ and f_X denote the corresponding cumulative distribution function, survival function and probability density function, respectively.
- $X =^{st} Y$ indicates that the random variables X and Y are identically distributed.
- A function f on a interval I is log convex (concave) if $f(\alpha x + (1-\alpha)y) \leqslant (\geqslant) f(x)^{\alpha} f(y)^{1-\alpha}$, $x, y \in I$, $0 \leqslant \alpha \leqslant 1$.
- A sequence a_n, $n = 0, 1, \ldots$ is log convex (concave) if $a_{n+1}^2 \leqslant (\geqslant) a_n a_{n+2}$.

A random variable X is

- increasing failure rate (IFR) if $\bar{F}_X$ is a log concave function.
- decreasing failure rate (DFR) if $\bar{F}_X$ is a log convex function.
- log concave if f_X is a log concave function.

The following properties apply

- If X log concave, then X is IFR.
- A non negative integer valued random variable L is IFR (DFR) if $\frac{P(L=n)}{P(L \geqslant n)}$ is increasing (decreasing) in n or equivalently $P(L \geqslant n)$ is a log concave (convex) sequence.

X is less or equal than Y in

- usual stochastic order, $X \leqslant_{st} Y$, if $F_X(x) \geqslant F_Y(x)$ or equivalently $E[f(X)] \leqslant E[f(Y)]$ for all f increasing function.
- hazard rate stochastic order, $X \leqslant_{hr} Y$, if $\frac{\bar{F}_Y(x)}{\bar{F}_X(x)}$ is increasing in x.
- reversed hazard rate stochastic order, $X \leqslant_{rhr} Y$, if $\frac{F_Y(x)}{F_X(x)}$ is increasing in x.
- likelihood ratio stochastic order, $X \leqslant_{lr} Y$, if $\frac{f_Y(x)}{f_X(x)}$ is increasing in x.
- In addition the following bivariate functional characterization of the likelihood ratio order can be found in Shanthikumar and Yao[23]: let X^* and Y^* be independent random variables

verifying $X =^{st} X^*$ and $Y =^{st} Y^*$ and H a bivariate function such that $H(x, y) - H(y, x) \geqslant 0$ for $x \leqslant y$, then $X \leqslant_{lr} Y$ if and only if $E[H(X^*, Y^*)] \geqslant E[H(Y^*, X^*)]$.

The following chain of implications concerning stochastic orders holds

$$\begin{array}{ccccc} & & hr & & \\ & \nearrow & & \searrow & \\ lr & & & & st \\ & \searrow & & \nearrow & \\ & & rhr & & \end{array}$$

Appendix A2. Auxiliary results

Let S_M^1 be a gamma random variable with shape parameter M and scale parameter 1.

Lemma 1. *$S_M^1 \leqslant_* S_{M+1}^1$, with $* =$ st, hr, rhr, lr.*

Proof. It follows that

$$\frac{f_{S_{M+1}^1}(x)}{f_{S_M^1}(x)} = \frac{\frac{1}{M!} x^M e^{-x}}{\frac{1}{(M-1)!} x^{M-1} e^{-x}} = \frac{x}{M}$$

is increasing in x, thus $S_M^1 \leqslant_{lr} S_{M+1}^1$ which implies the st, hr and rhr stochastic orders.

Lemma 2. Let f, g non null positive functions such that $\frac{f}{g}$ is decreasing. If $X \leqslant_{lr} Y$, then

$$E[f(X)]E[g(Y)] - E[f(Y)]E[g(X)] \geqslant 0.$$

Proof. If $\frac{f}{g}$ is decreasing, then $H(x, y) = f(x)g(y)$ verifies the condition for the bivariate characterization of likelihood ratio stochastic order in Shanthikumar and Yao[23] and the result follows.

Next, we define some auxiliary functions:

$$G(x) = \int_x^\infty \frac{1}{y} \pi\left(\frac{y}{\Lambda(T)}\right) dy.$$

For $M = 1, 2, \ldots$ and $0 \leqslant u \leqslant 1$:

$$\begin{aligned} A_M(u) &= \int_0^\infty \frac{1}{z} F_{S_M^1}(\Lambda(T)zu)\pi(z)dz. \\ A_M(1) &= \int_0^\infty \frac{1}{z} F_{S_M^1}(\Lambda(T)z)\pi(z)dz \\ &= \int_0^\infty f_{S_M^1}(x) \int_x^\infty \frac{1}{y} \pi\left(\frac{y}{\Lambda(T)}\right) dy dx \\ &= E[G(S_M^1)]. \end{aligned}$$

$$\begin{aligned} B_M(u) &= \int_0^\infty F_{S_M^1}(\Lambda(T)zu)\pi(z)dz \\ &= E[F_{S_M^1}(\Lambda(T)uZ)]. \\ B_M(1) &= \int_0^\infty f_{S_M^1}(x) \int_x^\infty \frac{1}{\Lambda(T)} \pi\left(\frac{z}{\Lambda(T)}\right) dz dx \\ &= E\left[\bar{F}_Z\left(\frac{S_M^1}{\Lambda(T)}\right)\right]. \end{aligned}$$

The previous identity follows since

$$F_{S_M^1}(\Lambda(T)z) = P\left(S_M^1 \leqslant \Lambda(T)z\right) = P\left(\frac{S_M^1}{\Lambda(T)z} \leqslant 1\right).$$

Lemma 3. *$\frac{A_M(1)}{B_M(1)}$ is decreasing in M.*

Proof. $\frac{A_M(1)}{B_M(1)}$ is decreasing in M if

$$\begin{aligned} &\left(E[G(S_M^1)]E\left[\bar{F}_Z\left(\frac{S_{M+1}^1}{\Lambda(T)}\right)\right]\right. \\ &\qquad \left. -E[G(S_{M+1}^1)]E\left[\bar{F}_Z\left(\frac{S_M^1}{\Lambda(T)}\right)\right]\right) \geqslant 0. \end{aligned} \tag{11}$$

Next, we verify that Conditions in Lemma 2 hold:

1. $H(x) = \frac{G(x)}{\bar{F}_Z\left(\frac{x}{\Lambda(T)}\right)}$ is decreasing in x since the sign of $\frac{dH(x)}{dx}$ is that of the following expression

$$\begin{aligned} &-\frac{1}{x}\pi\left(\frac{x}{\Lambda(T)}\right)\bar{F}_Z\left(\frac{x}{\Lambda(T)}\right) \\ &\quad + \frac{1}{\Lambda(T)}\pi\left(\frac{x}{\Lambda(T)}\right)\int_x^\infty \frac{1}{z}\pi\left(\frac{z}{\Lambda(T)}\right)dz \leqslant 0 \end{aligned}$$

2. $S_M^1 \leqslant_{lr} S_{M+1}^1$ from Lemma 1.

Thus, the inequality in equation (11) is derived from Lemma 2.

Lemma 4. *Let X_M and Y_M be random variables with probability density functions f_{X_M} and f_{Y_M} given, respectively, as follows*

$$f_{X_M}(u) = \frac{1}{A_M(1)} \int_0^\infty \frac{1}{z} f_{\frac{S_M^1}{\Lambda(T)z}}(u)\pi(z)dz, \quad 0 \leqslant u \leqslant 1$$

and

$$f_{Y_M}(u) = \frac{1}{B_M(1)} \int_0^\infty f_{\frac{S_M^1}{\Lambda(T)z}}(u)\pi(z)dz, \quad 0 \leqslant u \leqslant 1$$

If $\frac{\pi(ax)}{\pi(x)}$ is an increasing function in x for $0 \leqslant a \leqslant 1$, then $X_M \leqslant_{st} X_{M+1}$ and $Y_M \leqslant_{st} Y_{M+1}$.

Proof. $\frac{\pi(ax)}{\pi(x)}$ increasing in x for $0 \leqslant a \leqslant 1$, implies that $u\Lambda(T)Z \leqslant_{lr} \Lambda(T)Z$ which is one of the assumptions in Lemma 2.

Since the following equality applies

$$\begin{aligned} F_{X_M}(u) &= \frac{A_M(u)}{A_M(1)} \\ &= E\left[\frac{1}{Z}F_{S^1_M}(\Lambda(T)uZ)\right] / E\left[\frac{1}{Z}F_{S^1_M}(\Lambda(T)Z)\right] \end{aligned}$$

then, $F_{X_M}(u) \geqslant F_{X_{M+1}}(u)$ is equivalent to

$$\begin{aligned}\Big(&E\left[\tfrac{1}{Z}F_{S^1_M}(\Lambda(T)uZ)\right]E\left[\tfrac{1}{Z}F_{S^1_{M+1}}(\Lambda(T)Z)\right] \\ &-E\left[\tfrac{1}{Z}F_{S^1_{M+1}}(\Lambda(T)uZ)\right]E\left[\tfrac{1}{Z}F_{S^1_M}(\Lambda(T)Z)\right]\Big) \geqslant 0\end{aligned}$$

From Lemma 1 and the relation between the likelihood and usual stochastic orders, it follows that $\frac{\frac{1}{x}F_{S^1_M}(x)}{\frac{1}{x}F_{S^1_{M+1}}(x)}$ is decreasing. Thus the previous inequality is derived from Lemma 2.

Likewise, $F_{Y_M}(u) = \frac{B_M(u)}{B_M(1)}$, thus $F_{Y_M}(u) \geqslant F_{Y_{M+1}}(u)$ iff

$$\begin{aligned}\Big(&E[F_{S^1_M}(\Lambda(T)uZ)]E[F_{S^1_{M+1}}(\Lambda(T)Z)] \\ &-E[F_{S^1_{M+1}}(\Lambda(T)uZ)]E[F_{S^1_M}(\Lambda(T)Z)]\Big) \geqslant 0\end{aligned}$$

From Lemma 1, $\frac{F_{S^1_M}(x)}{F_{S^1_{M+1}}(x)}$ is decreasing and conditions in Lemma 2 apply.

Note: Observe that $f_{\frac{S^1_M}{\Lambda(T)z}}(u) = f_{S^1_M}(\Lambda(T)zu)\Lambda(T)z$.

Remark 1. The following alternative expression for $A_M(1)$ and $B_M(1)$ also apply:

$$\begin{aligned} A_M(1) &= \int_0^1 \int_0^\infty \frac{1}{z} f_{\frac{S^1_M}{\Lambda(T)z}}(u)\pi(z)dzdu \\ B_M(1) &= \int_0^1 \int_0^\infty f_{\frac{S^1_M}{\Lambda(T)z}}(u)\pi(z)dzdu \end{aligned}$$

Lemma 5. *Let X be a random variable, then:*

(a) *$E[f(X)g(X)] \geqslant E[f(X)]E[g(X)]$ if f and g are both increasing or decreasing functions.*

(b) *$E[f(X)g(X)] \leqslant E[f(X)]E[g(X)]$ if f is increasing and g is decreasing.*

The proof of this result can be found in Cuadras[24] and Joe.[25]

Lemma 6. *Let $\{N_1(t) : t \geqslant 0\}$ be a homogeneous Poisson process with failure rate $\lambda(t) = 1$ and X a non negative random variable independent from the process and density function f_X, it follows that*

(a) *$N_1(X)$ is IFR (DFR, log concave) if X is IFR (DFR, log concave).*

(b) *Let U be a uniform random variable on the interval $[0,1]$ independent from both the process and X. In addition g is an increasing function in $[0,T]$, then $N_1\left(\frac{g(UT)}{g(T)}X\right) \leqslant_{lr} N_1(X)$ if $\frac{f_X(ax)}{f_X(x)}$ is increasing in x for $0 \leqslant a \leqslant 1$.*

Proof. The preservation of the ageing classes given in *(a)* are well known results (see Grandell[26]).

Case (b) is based on the following property stated in Shaked and Shanthikumar[27]: If $A \leqslant_{lr} B$, then $N_1(A) \leqslant_{lr} N_1(B)$. Let $A = \frac{g(UT)}{g(T)}X$ and $B = X$ since the following identity applies

$$\frac{f_A(x)}{f_B(x)} = \frac{1}{f_X(x)}\int_0^1 \frac{g(T)}{g(uT)} f_X\left(\frac{xg(T)}{g(uT)}\right)du$$

$f_X\left(\frac{xg(T)}{g(uT)}\right)/f_X(x)$ is decreasing in x and so is $\frac{f_A(x)}{f_B(x)}$, therefore the conclusion holds.

Remark 2. Consider the mixed NHPP repair process $\{N(t) : t \geqslant 0\}$. Observe that $N(T) =^{st} N_1(\Lambda(T)Z)$:

$$\begin{aligned} &P(N_1(\Lambda(T)Z) = M) \\ &= \int_0^\infty P(N_1(\Lambda(T)z) = M)\pi(z)dz \\ &= \int_0^\infty e^{-\Lambda(T)z}\frac{(\Lambda(T)z)^M}{(M-1)!}\pi(z)dz = P(N(T) = M). \end{aligned}$$

Moreover, we have that $B_M(1) = P(N(T) \geqslant M)$:

$$\begin{aligned} P(N(T) \geqslant M)) &= P(N_1(\Lambda(T)Z) \geqslant M) \\ &= \int_0^\infty P(S^1_M \leqslant \Lambda(T)z)\pi(z)dz \\ &= \int_0^\infty \pi(z)\int_0^{\Lambda(T)z} f_{S^1_M}(x)dxdz \\ &= \int_0^\infty f_{S^1_M}(x)\int_{x/\Lambda(T)}^\infty \pi(z)dzdx \\ &= \int_0^\infty f_{S^1_M}(x)\int_x^\infty \pi\left(\frac{u}{\Lambda(T)}\right)\frac{1}{\lambda(T)}dudx \end{aligned}$$

The change $\Lambda(T)z = u$ leads to the last identity in the foregoing expression.

The identical distribution of $N(T)$ and $N_1(\Lambda(T)Z)$ is used along with Lemma 6 (a) to derive the ageing properties of $N(T)$.

Remark 3. The use of Lemma 6 (b) is based on the following identity:

$$\int_0^T P(N(t)=M)dt = \int_0^T P\left(N_1\left(\frac{\Lambda(t)}{\Lambda(T)}\Lambda(T)Z\right)=M\right)dt = TP\left(N_1\left(\frac{\Lambda(UT)}{\Lambda(T)}\Lambda(T)Z\right)=M\right)$$

where U is a uniform random variable on the interval $[0,1]$ independent from the Poisson process.

Lemma 7.

$$\int_0^T P(N(t)=M)dt = A_{M+1}(1)E\left[\frac{1}{\lambda(H^{-1}(X_{M+1}))}\right]$$

H^{-1} is the inverse function of $H(t)=\frac{\Lambda(t)}{\Lambda(T)}$, $0\leqslant t\leqslant T$.

Proof. $N(T)={}^{\mathrm{st}}N_1(\Lambda(T)Z)$ and the change $u=\frac{\Lambda(t)}{\Lambda(T)}$ lead to

$$\int_0^T P(N(t)=M)dt = \int_0^T\int_0^\infty e^{-\Lambda(t)z}\frac{(\Lambda(t)z)^M}{M!}\pi(z)dzdt = \int_0^1 \frac{\Lambda(T)}{\lambda(H^{-1}(u))}\int_0^\infty e^{-\Lambda(T)zu}\frac{(\Lambda(T)zu)^M}{M!}\pi(z)dzdu = \int_0^1 \frac{\Lambda(T)}{\lambda(H^{-1}(u))}\int_0^\infty f_{S^1_{M+1}}(\Lambda(T)zu)\pi(z)dzdu = \int_0^1 \frac{1}{\lambda(H^{-1}(u))}\int_0^\infty \frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu.$$

Calculations in the following Lemma involve the numerator of the cost function in scenario 2, denoted as C(T,M). Thus, $C(T,M)=E[C(\tau)]$ with $E[C(\tau)]$ in equation (7).

Lemma 8. *If $c_a=c_f$, it follows that*

$$C(T,M+1)-C(T,M) = \int_0^1 c_1\left(H^{-1}(x)\right)\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(x)\pi(z)dzdx.$$

Proof:

$$C(T,M+1)-C(T,M) = MP(N(T)=M)\int_0^T c_1(x)\frac{\lambda(x)}{\Lambda(T)}dx + \sum_{k=M+1}^{\infty}\sum_{j=0}^{M-1}P(N(T)=k)k\binom{k-1}{j}\times \int_0^T c_1(x)\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^j\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-1-j}\frac{\lambda(x)}{\Lambda(T)}dx - \sum_{k=M}^{\infty}\sum_{j=0}^{M-2}P(N(T)=k)k\binom{k-1}{j}\times \int_0^T c_1(x)\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^j\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-1-j}\frac{\lambda(x)}{\Lambda(T)}dx \int_0^T c_1(x)\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^j\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-1-j}\frac{\lambda(x)}{\Lambda(T)}dx + \sum_{k=M+1}^{\infty}P(N(T)=k)k\binom{k-1}{M-1}\int_0^T c_1(x)\times \left(\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^{M-1}\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-1-M+1}\frac{\lambda(x)}{\Lambda(T)}\right)dx - MP(N(T)=M)\sum_{j=0}^{M-2}\binom{M-1}{j}\times \int_0^T c_1(x)\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^j\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{M-1-j}\frac{\lambda(x)}{\Lambda(T)}dx$$

$$= \sum_{k=M}^{\infty}P(N(T)=k)k\binom{k-1}{M-1}\int_0^T c_1(x)\times \left(\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^{M-1}\left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-1-M+1}\frac{\lambda(x)}{\Lambda(T)}\right)dx$$

$$= \int_0^T kc_1(x)\sum_{k=M}^{\infty}\int_0^\infty \frac{(z\Lambda(T))^k}{k!}e^{-z\Lambda(T)}\pi(z)\times \frac{(k-1)!}{(M-1)!(k-M)!}\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^{M-1}\times \left(1-\frac{\Lambda(x)}{\Lambda(T)}\right)^{k-M}\frac{\lambda(x)}{\Lambda(T)}dzdx$$

$$= \int_0^T c_1(x)\int_0^\infty \frac{1}{(M-1)!}\left(\frac{\Lambda(x)}{\Lambda(T)}\right)^{M-1}(z\Lambda(T))^M\times \left(e^{-\Lambda(T)z}e^{\Lambda(T)z(1-\frac{\Lambda(x)}{\Lambda(T)})}\frac{\lambda(x)}{\Lambda(T)}\pi(z)\right)dzdx$$

$$= \int_0^T c_1(x)\times \left(\int_0^\infty \frac{z}{(M-1)!}e^{-\Lambda(x)z}(\Lambda(x)z)^{M-1}\lambda(x)\pi(z)dz\right)dx$$

$$\int_0^T c_1(x)\int_0^\infty f_{S_M^1}(\Lambda(x)z)z\lambda(x)\pi(z)dzdx$$
$$\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty f_{S_M^1}(\Lambda(T)zu)z\Lambda(T)\pi(z)dzdu$$
$$\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty f_{\frac{S_M^1}{\Lambda(T)z}}(u)\pi(z)dzdu.$$

The first expression in the last line is obtained after the change $u = \frac{\Lambda(x)}{\Lambda(T)} = H(x)$.

Appendix A3. Existence of optimal policies in scenarios 1 and 2

Consider the following function

$$L(T, M) = C(T, M)\frac{\int_0^T P(N(t) = M)dt}{C(T, M+1) - C(T, M)} - \sum_{j=0}^{M-1}\int_0^T P(N(t) = j)dt.$$

After basic calculations we obtain

$$Q(T, M+1) \leqslant Q(T, M) \Leftrightarrow L(T, M) \geqslant 0,\ M = 1, 2, \ldots \tag{12}$$

If the following condition applies

$$\lim_{M\to\infty}\frac{\int_0^T P(N(t) = M)dt}{C(T, M+1) - C(T, M)} = 0$$

then

$$\lim_{M\to\infty} L(T, M) = -T < 0 \tag{13}$$

$M \to \infty$ implies that only the age replacement at T occurs and, from the univariate policies in Section 4, we have

$$\lim_{M\to\infty} C(T, M) = E(Z)\Lambda(T)\int_0^T c_1(t)\frac{\lambda(t)}{\Lambda(T)}dt + c_a < \infty.$$

In addition

$$L(T, M) - L(T, M+1) = C(T, M+1) \times \left(\frac{\int_0^T (P(N(t) = M)dt}{C(T, M+1) - C(T, M)} - \frac{\int_0^T (P(N(t) = M+1)dt}{C(T, M+2) - C(T, M+1)}\right)$$

and the next property follows:

$L(T, M)$ is decreasing in M iff

$$\frac{\int_0^T P(N(t) = M)dt}{C(T, M+1) - C(T, M)} \text{ is decreasing in } M. \tag{14}$$

Proof of Proposition 2. From equation (14)$L(T, M)$ is decreasing in M and, hence, $L(T, M) \leqslant L(T, 1)$. In case that $L(T, 1) \leqslant 0$, it follows that $L(T, M) \leqslant 0$. From equation (12), $Q(T, M)$ is also decreasing in M and the conclusion holds for $M^* = \infty$.

Condition (9) leads to (13) and, if $L(T, 1) > 0$, there exits M such that $L(T, M) < 0$. Let $M_0 = \inf\{M > 1 : L(T, M) < 0\}$. Since $L(T, M)$ is decreasing in M, it follows that $L(T, M) \geqslant 0$ for $M = 1, \ldots, M_0$ and $L(T, M) < 0$ for $M = M_0, M_0 + 1, \ldots$. Then, from equation (12) it follows that $Q(T, M)$ is decreasing for $M = 1, \ldots, M_0$ and increasing for $M = M_0, M_0 + 1, \ldots$ and the result holds with $M^\star = M_0$.

Proof of Theorem 1. ***a***) (Scenario 1). The expression of $C(T, M) = E[C(\tau)]$ in equation (6) for scenario 1, leads to

$$\begin{aligned}&C(T, M+1) - C(T, M)\\ &= c_1 P(N(T) \geqslant M) + \left(c_a - c_f\right)P(N(T) = M)\\ &= c_1 P(N(T) \geqslant M+1) + \left(c_1 + c_a - c_f\right)P(N(T) = M)\end{aligned}$$

In addition, according to Remark 2, the following identity holds

$$P(N(T) \geqslant M+1) = B_{M+1}(1)$$

and, considering Lemma 7, the function in equation (8) can be also expressed as follows

$$\frac{\int_0^T P(N(t) = M)dt}{C(T, M+1) - C(T, M)} = \frac{A_{M+1}(1)}{B_{M+1}(1)}\frac{1}{\left(c_1 + \left(c_1 + c_a - c_f\right)\frac{P(N(T) = M)}{P(N(T) \geqslant M+1)}\right)} \times E\left[\frac{1}{\lambda(H^{-1}(X_{M+1}))}\right] \tag{15}$$

Next, we prove that if a_1) holds, then assumptions in Proposition 2 for scenario 1 apply. First, we obtain that equation (8) is decreasing in M:

1. $\frac{A_{M+1}(1)}{B_{M+1}(1)}$ is decreasing in M (Lemma 3).
2. Z being log concave implies that $N_1(Z)$ is also log concave (Lemma 6 (a)) and so is $N_1(\Lambda(T)Z)$ since $\Lambda(T)$ is a constant. From Remark 2) the log concavity of $N(T)$ is derived and, therefore,

$N(T)$ is IFR. Thus, $\frac{P(N(T)=M)}{P(N(T)\geqslant M+1)}$ is increasing in M and $(c_1+c_a-c_f)\frac{P(N(T)=M)}{P(N(T)\geqslant M+1)}$ is also increasing given condition $c_1+c_a-c_f>0$ in Theorem 1.

3. From Lemma 4, the random variable X_{M+1} that takes values in $[0,1]$ with density function $\int_0^\infty \frac{1}{A_{M+1}(1)}\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dz$, is increasing in M in the usual stochastic order.

Hence,

$$E\left[\frac{1}{\lambda(H^{-1}(X_{M+1}))}\right]=\int_0^1\frac{1}{\lambda(H^{-1}(u))}\times\left(\int_0^\infty\frac{1}{A_{M+1}(1)}\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dz\right)du$$

is decreasing in M since $\lambda(x)$ is assumed to be increasing under conditions a_1) and a_2).

The following alternative expression for equation (8) leads to prove that this function is decreasing in M under conditions a_2) in Theorem 1.

$$\frac{\int_0^T P(N(t)=M)dt}{C(T,M+1)-C(T,M)}=\frac{\int_0^T P(N(t)=M)dt}{P(N(T)=M)}\frac{1}{\frac{c_1P(N(T)\geqslant M)}{P(N(T)=M)}+c_a-c_f}. \tag{16}$$

1. If Z is DFR, Lemma 6 (a) and Remark 2 imply that $N(T)$ is DFR, that is, $\frac{P(N(T)=M)}{P(N(T)\geqslant M)}$ is decreasing in M.
2. Both, Remarks 2 and 3 lead to the following identity:

$$\frac{\int_0^T P(N(t)=M)dt}{P(N(T)=M)}=\frac{TP\left(N_1\left(\frac{\Lambda(UT)}{\Lambda(T)}\Lambda(T)Z\right)=M\right)}{P(N_1(\Lambda(T)Z)=M)}.$$

Under assumptions in a_2), $\frac{\pi(az)}{\pi(z)}$ is increasing in z for $0\leqslant a\leqslant 1$, then from Lemma 6 (b), it follows that $N_1\left(\frac{\Lambda(UT)}{\Lambda(T)}\Lambda(T)Z\right)\leqslant_{lr}N_1(\Lambda(T)Z)$. Hence, $\frac{\int_0^T P(N(t)=M)dt}{P(N(T)=M)}$ is decreasing in M and so is equation (16).

Next, we prove that the limiting condition (9) in Proposition 2 holds under assumptions in Theorem 1. Condition $c_1+c_a-c_f>0$, the stochastic order in X_M and $\frac{1}{\lambda(x)}$ being a decreasing function apply under assumptions a_1) and a_2), leading to the following inequality

$$\frac{\int_0^T P(N(t)=M)}{C(T,M+1)-C(T,M)}\leqslant\frac{A_{M+1}(1)}{B_{M+1}(1)}\times\frac{1}{\left(c_1+(c_a+c_1-c_f)\frac{P(N(T)=M)}{P(N(T)\geqslant M+1)}\right)}E\left[\frac{1}{\lambda(H^{-1}(X_1))}\right]\leqslant\frac{A_{M+1}(1)}{c_1B_{M+1}(1)}E\left[\frac{1}{\lambda(H^{-1}(X_1))}\right].$$

In addition, according to Remark 1, it follows that

$$\frac{A_{M+1}(1)}{B_M(1)}=\frac{\int_0^1\int_0^\infty\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu}{\int_0^1\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(u)\pi(z)dzdu}\leqslant\frac{\Lambda(T)}{M}. \tag{17}$$

The last inequality applies since

$$\begin{aligned}&\int_0^1\int_0^\infty\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu\\&=\int_0^1\int_0^\infty\frac{1}{z}f_{S^1_{M+1}}(\Lambda(T)zu)\Lambda(T)z\pi(z)dzdu\\&=\int_0^1\int_0^\infty\frac{1}{z}\frac{1}{M!}e^{-\Lambda(T)zu}\Lambda(T)^M(zu)^M\Lambda(T)z\pi(z)dzdu\\&=\frac{\Lambda(T)}{M}\int_0^1\frac{1}{(M-1)!}u\times\left(\int_0^\infty e^{-\Lambda(T)zu}\Lambda(T)^{M-1}(zu)^{M-1}\Lambda(T)z\pi(z)dz\right)du\\&=\frac{\Lambda(T)}{M}\int_0^1 u\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(u)\pi(z)dzdu\\&\leqslant\frac{\Lambda(T)}{M}\int_0^1\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(u)\pi(z)dzdu.\end{aligned}$$

In addition, the following function derived from Remark 2, can be expressed in the two next alternative ways that serve to prove that it is bounded either under conditions a_1) or a_2).

$$\frac{B_M(1)}{B_{M+1}(1)}=1+\frac{P(N(T)=M)}{P(N(T)\geqslant M+1)} \tag{18}$$

$$=1+\frac{P(N(T)=M)}{P(N(T)=M+1)}\frac{P(N(T)=M+1)}{P(N(T)\geqslant M+1)}=\frac{P(N(T)\geqslant M)}{P(N(T)\geqslant M+1)}. \tag{19}$$

For Z log concave (condition in a_1)), $\frac{P(N(T)=M)}{P(N(T)=M+1)}$ in equation (18) is decreasing in M by Lemma 6 (a) and Remark 2.

For Z DFR (condition in a_2)), $N(T)$ is DFR by Lemma 6 (a) and Remark 2. Hence, $\frac{P(N(T)\geqslant M)}{P(N(T)\geqslant M+1)}$ in equation (19) is decreasing in M.

Therefore, applying equation (17), the product $\frac{A_{M+1}(1)}{B_M(1)}\frac{B_M(1)}{B_{M+1}(1)} = \frac{A_{M+1}(1)}{B_{M+1}(1)}$ tends to zero when M tends to infinity and, thus, equation (9) applies. Therefore, conditions in Proposition 2 are fulfilled either a_1) or a_2) holds.

Proof of Theorem 1 *b*) (Scenario 2). Now we consider scenario 2 with $c_a = c_f$. Applying Lemma 8 we have

$$\begin{aligned} &C(T,M+1)-C(T,M)\\ &=\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(u)\pi(z)dzdu\\ &=B_M(1)E\left[c_1\left(H^{-1}(Y_M)\right)\right]. \end{aligned}$$

The previous identity and Lemma 7 lead to

$$\frac{\int_0^T P(N(t)=M)}{C(T,M+1)-C(T,M)} = \frac{A_{M+1}(1)}{B_M(1)}\frac{E\left[\frac{1}{\lambda(H^{-1}(X_{M+1}))}\right]}{E[c_1(H^{-1}(Y_M))]}. \tag{20}$$

The following properties hold:

1. $\frac{A_{M+1}(1)}{B_{M+1}(1)}$ is decreasing in M by Lemma 3. $\frac{B_{M+1}(1)}{B_M(1)}$ is decreasing in M by Lemma 6 (a) and Remark 2 since Z is IFR. Therefore, $\frac{A_{M+1}(1)}{B_M(1)}$ is also decreasing in M.
2. Since $\frac{\pi(az)}{\pi(z)}$ is increasing in z for $0\leqslant a\leqslant 1$, and both $c_1(x)$ and $\lambda(x)$ increasing functions in x, then $E\left[\frac{1}{\lambda(H^{-1}(X_{M+1}))}\right]$ and $\frac{1}{E[c_1(H^{-1}(Y_M))]}$ are decreasing in M by Lemma 4.

Therefore, under conditions given in b) for scenario 2, the result in equation (14) follows.

Next, we prove that equation (9) also applies:

$$\begin{aligned} &\int_0^T P(N(t)=M)dt\\ &=\int_0^T\int_0^\infty e^{-\Lambda(t)z}\frac{(\Lambda(t)z)^M}{M!}\pi(z)dzdt\\ &=\int_0^1\frac{1}{\lambda(H^{-1}(u))}\int_0^\infty\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu\\ &=A_{M+1}(1)\int_0^1\frac{1}{\lambda(H^{-1}(u))}\times\\ &\left(\int_0^\infty\frac{1}{A_{M+1}(1)z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dz\right)du\\ &\leqslant A_{M+1}(1)\int_0^1\frac{1}{c_1(H^{-1}(u))\lambda(H^{-1}(u))}\times\\ &\int_0^\infty\frac{1}{A_{M+1}(1)z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu\\ &\times\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty\frac{1}{A_{M+1}(1)z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu\\ &=E\left[\frac{1}{c_1(H^{-1}(X_M))\lambda(H^{-1}(X_M))}\right]\times\\ &\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu. \end{aligned}$$

Regarding the foregoing expressions, the last one in the first line is obtained after the change of variable $u=\frac{\Lambda(t)}{\Lambda(T)}=H(t)$ in the integral. Thus, $t=H^{-1}(u)$ and $dt=\frac{\lambda(T)}{\lambda(t)}du$.

The inequality in the third line follows applying Lemma 5 (b), since $c_1(x)$ is increasing and $\frac{1}{c_1(x)\lambda(x)}$ is decreasing. Furthermore

$$\begin{aligned} &\int_0^1 c_1\left(H^{-1}(u)\right)\int_0^\infty\frac{1}{z}f_{S^1_{\frac{M+1}{\Lambda(T)z}}}(u)\pi(z)dzdu\\ &=\frac{\Lambda(T)}{M}\int_0^1 c_1\left(H^{-1}(u)\right)\times\\ &\left(\int_0^\infty\frac{1}{(M-1)!}\Lambda(T)^M(zu)^M e^{-\Lambda(T)zu}\pi(z)dz\right)du\\ &\leqslant\frac{\Lambda(T)}{M}\int_0^1 c_1\left(H^{-1}(u)\right)\frac{1}{(M-1)!}\times\\ &\left(\int_0^\infty\Lambda(T)^{M-1}z^{M-1}u^{M-1}e^{-\Lambda(T)zu}\lambda(T)z\pi(z)dz\right)du\\ &=\frac{\Lambda(T)}{M}\int_0^1 c_1\left(H^{-1}(u)\right)\left(\int_0^\infty f_{S^1_{\frac{M}{\Lambda(T)z}}}(u)\pi(z)dz\right)du\\ &=\frac{\Lambda(T)}{M}B_M(1)E\left[c_1\left(H^{-1}(Y_M)\right)\right]. \end{aligned}$$

Lemma 8 and the previous inequality lead to

$$\begin{aligned} &\frac{\int_0^T P(N(t)=M)}{C(T,M+1)-C(T,M)}\\ &\leqslant\frac{\Lambda(T)}{M}E\left[\frac{1}{c_1(H^{-1}(X_M))\lambda(H^{-1}(X_M))}\right]\\ &\leqslant\frac{\Lambda(T)}{M}E\left[\frac{1}{c_1(H^{-1}(X_1))\lambda(H^{-1}(X_1))}\right]. \end{aligned}$$

The last inequality follows since the assumption $c_1(x)$ and $\lambda(x)$ being increasing functions in Theorem 1 *b*), implies that $\frac{1}{c_1(H^{-1}(u))}$, $\frac{1}{\lambda(H^{-1}(u))}$ are both decreasing. In addition, $X_1 \leqslant_{st} X_M$ follows from the assumption $\frac{\pi(az)}{\pi(z)}$, $0 \leqslant a \leqslant 1$ in Theorem 1 *b*) and Lemma 4.

Hence, when M tends to infinity, condition (9) holds. Thus, Proposition 2 applies under assumptions of scenario 2 in Theorem 1 *b*).