Maintaining Systems with Heterogeneous Spare Parts

David T. Abdul-Malak[1]
Department of Industrial Engineering
University of Pittsburgh
1025 Benedum Hall
3700 O'Hara Street
Pittsburgh, PA 15261 USA

Jeffrey P. Kharoufeh[2]
Department of Industrial Engineering
Clemson University
100 Freeman Hall
Clemson, SC 29634 USA

Lisa M. Maillart[3]
Department of Industrial Engineering
University of Pittsburgh
1025 Benedum Hall
3700 O'Hara Street
Pittsburgh, PA 15261 USA

Accepted Version
July 3, 2019

Abstract

We consider the problem of optimally maintaining a stochastically degrading, single-unit system using heterogeneous spares of varying quality. The system's failures are unannounced; therefore, it is inspected periodically to determine its status (functioning or failed). The system continues in operation until it is either preventively or correctively maintained. The available maintenance options include perfect repair, which restores the system to an as-good-as-new condition, and replacement with a randomly-selected unit from the supply of heterogeneous spares. The objective is to minimize the total expected discounted maintenance costs over an infinite time horizon. We formulate the problem using a mixed observability Markov decision process (MOMDP) model in which the system's age is observable but its quality must be inferred. We show, under suitable conditions, the monotonicity of the optimal value function in the belief about the system quality and establish conditions under which finite preventive maintenance thresholds exist. A detailed computational study reveals that the optimal policy encourages exploration when the system's quality is uncertain; the policy is more exploitive when the quality is highly certain. The study also demonstrates that substantial cost savings are achieved by utilizing our MOMDP-based method as compared to more naïve methods of accounting for heterogeneous spares.

[1]Email: dta10@pitt.edu.
[2]Corresponding author. Ph: (864) 656-5540; Email: kharouf@clemson.edu.
[3]Email: maillart@pitt.edu.

1 Introduction

Operations and maintenance activities constitute a significant proportion of the operational budgets of many organizations [6]. Consequently, companies are employing increasingly sophisticated methods to improve system reliability, ensure safety, and reduce costs. A common assumption in maintenance models is that available spares (components or sub-systems) originate from a homogeneous population in which each spare has identical degradation characteristics. However, in reality, replacement parts may exhibit significant unit-to-unit variability in their quality characteristics. For instance, micro-electro-mechanical systems (MEMS) are known to suffer from a variety of local defects including particulate, ionic, organic, and isolated defects (e.g., voids and stringers) that can cause substantial unit-to-unit variability [9]. More generally, multiple device qualities can stem from manufacturing processes that are still in early stages of development and, therefore, highly variable [31]. Irrespective of the source of heterogeneity, ignoring this variability can reduce overall system reliability and lead to economic losses [4].

In this paper, we consider the problem of optimally maintaining a stochastically degrading, single-unit system with heterogeneous spare parts. Specifically, the spare parts originate from Y distinct and heterogeneous subpopulations, each of which has its own time-to-failure distribution. Failures are not self-announcing; therefore, the system is inspected periodically to determine its status (functioning or failed). The system continues in operation until it is either preventively or correctively maintained. The available maintenance options include perfect repair, which restores the system to an as-good-as-new condition, or replacement of the system with a randomly-selected unit from the lot of heterogeneous spares. The primary advantage of this framework is that over time, as a system remains in operation, the subpopulation from which it originates becomes more apparent. Consequently, we are able to make better-informed maintenance decisions, thereby reducing long-run maintenance costs. In contrast to other models that consider spare part heterogeneity, we are able to update our beliefs, even in the absence of detailed condition monitoring information, based on the system's age alone. Assuming an intuitive cost structure that includes inspection, preventive maintenance, and corrective maintenance costs, our objective is to obtain a cost-minimizing policy that accounts for preventive and corrective maintenance decisions. To this end, we formulate an infinite-horizon, discounted mixed observability Markov decision process (MOMDP) model and establish important properties of the cost function and optimal policy.

Within the applied probability and operations research communities, maintenance optimization models have been studied extensively over the last five decades. Many existing surveys highlight the most prevalent models for both single- and multi-component systems [1, 6, 13, 18, 19, 21, 22, 23, 25, 27]. Particularly relevant to our work here is the survey of Valdez-Flores and Feldman [25], which reviews the maintenance optimization literature of single-unit systems and includes models for inspection, minimal repair, shock, and replacement. Over the past two decades, reliability and maintenance models have increasingly addressed the problem of population heterogeneity. A common method for handling heterogeneous populations is to eliminate low-quality subpopulations before they enter field service through burn-in procedures. Burn-in procedures are tests engineered to stress and detect devices that are likely to incur early failures (infant mortality). For a general background on burn-in procedures and models, the reader is referred to [10, 11] and references contained therein. Mi [14, 15] explores the joint problems of burn-in, maintenance, and repair when a system exhibits a bathtub-shaped failure rate function and characterizes the optimal burn-in times and maintenance policies. While early burn-in models focused on lifetime-based burn-in (discarding units that have failed before the end of the burn-in period), recent models are concerned with degradation-based burn-in (discarding units that have a particular degradation level at the end of the burn-in period). For example, Ye et al. [32] consider a joint burn-in and maintenance problem when the degradation processes of two different subpopulations are modeled as Weiner processes with distinct drift parameters. Optimal burn-in and replacement policies are characterized for age- and block-replacement and shown to be effective as compared to traditional lifetime-based burn-in approaches. Xiang et al. [30] investigate the more general case of joint burn-in and preventive replacement when there are n subpopulations subject to stochastic degradation. Their framework is extended to the case of burn-in under accelerated conditions (e.g., elevated voltage, humidity, and temperature) with condition-based maintenance (CBM).

In addition to burn-in methods, another fruitful research area related to heterogeneous populations is the modeling of degradation in the presence of unit-to-unit variability. A common strategy is to augment a standard degradation model by allowing for some (or all) of the model's parameters to be random (i.e., to differ between systems within the population). This strategy has been employed when the degradation is modeled as a Brownian motion process [2, 17, 29], a gamma process [12, 24], or when it assumes an exponential form [8]. For each of these model types, the aim is to provide some measure of remaining useful life. Although these frameworks find wide applicability in CBM applications, they typically require that the system deteriorates with time, the deterioration level is observable at any time, and the device fails when the degradation level reaches a specified threshold [28].

Some recent papers consider CBM strategies with heterogeneous populations. Chen et al. [5] consider a system whose degradation evolves as an inverse Gaussian process with random effects. Similarly, Elwany et al. [8] consider a random effects model with a system degrading according to an exponential form with sensors monitoring the system's degradation. For both models, conditions are provided under which monotone control-limit policies are optimal. In lieu of a true burn-in procedure or CBM strategy, Zhang et al. [33] propose a joint inspection-replacement policy. Each system is allowed to enter service, but an early inspection is conducted to determine its health state, at which point the unit is either replaced (if it appears to be defective) or a preventive replacement time is determined based on the health state. They show that this inspection-replacement policy outperforms a joint burn-in and age-based replacement policy. More recently, and most relevant to our work here, van Oosterom et al. [26] study a Markov deteriorating system (i.e., one with finitely many condition states) that originates from a population of heterogeneous spares. Assuming finitely many system qualities, they provide a set of conditions under which the optimal maintenance policy is a threshold-type policy. We extend the model in [26] by removing restrictions on the system's time-to-failure distribution and consider both repair and replacement actions.

Specifically, our problem setting is unique in that (i) we do not impose a finite-state Markovian structure on the degradation process; (ii) we cannot directly observe degradation (or a degradation signal); rather, we can only determine if the system is failed or working upon inspection; and (iii) we consider unit-to-unit variability in repairable systems. Within this framework, knowledge about the system's quality can be learned and leveraged without a well-defined notion of degradation or failure threshold, or even an observable degradation signal of any kind. That is, as the system continues in operation and undergoes repairs over time, we are able to glean valuable information about its population of origin based on its age alone. By way of Bayesian inference procedures, we update our understanding of the system's quality and make better-informed decisions that lead to demonstrable cost savings. Our main contributions include a novel MOMDP framework for accounting for spare part heterogeneity; providing conditions under which the optimal value function is monotone in the belief space; providing conditions under which the optimal policy calls for preventive maintenance; characterizing the optimal policy; and executing a computational study that demonstrates the utility of our proposed framework and provides additional insights into the optimal policy as an exploration/exploitation type policy.

The remainder of this paper is organized as follows. In Section 2, we present the maintenance problem when unit-to-unit variability is significant and formulate a mathematical model of the corresponding sequential decision-making process. Section 3 discusses attributes of the value function and the optimal policy. Finally, in Section 4, we provide a detailed computational study that illustrates the importance of accounting for heterogeneity, demonstrates the effectiveness of our modeling framework, and highlights some interesting properties of the optimal maintenance policy.

2 Model Formulation

Consider a single-unit, repairable system that begins operation in an as-good-as-new condition, that is, with an operating age of 0. The system continues functioning until it fails, i.e., when its operating age exceeds a probabilistically determined time-to-failure, or until a maintenance action is taken preventively. In what follows, all random variables are defined on a common, complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

It is assumed that the system originates from a lot of spare systems that, despite being visually indistinguishable, have distinct time-to-failure distributions. The lot of spare systems is composed of Y $(Y > 1)$ distinct qualities, and we denote the finite set of system types by $\mathcal{Y} := \{1, 2, \ldots, Y\}$. To ensure that the proportion of each type of system remains constant, we assume that the lot of spare systems is infinitely large. Additionally, we assume that the proportion of systems is known *a priori* and given by the vector $\boldsymbol{\rho} := (\rho_1, \rho_2, \ldots, \rho_Y)$, where ρ_y denotes the proportion of systems of quality $y \in \mathcal{Y}$ and

$$\sum_{y \in \mathcal{Y}} \rho_y = 1.$$

When a system is placed into service, it is able to operate until its (random) time-to-failure T. We denote the distribution function (d.f.) of T by $F(t) := \mathbb{P}(T \leq t)$, $t \geq 0$, and define $\bar{F}(t)$ as the complementary d.f., or survival function $\bar{F}(t) = 1 - F(t)$. Additionally, we assume that each quality has its own (known) failure distribution. We let Q denote the (random) quality of the system; hence, $\rho_y = \mathbb{P}(Q = y)$. Define $T_y = [T|Q = y]$ as the conditional time-to-failure, given $Q = y$, with distribution function $F_y(t) = \mathbb{P}(T \leq t|Q = y)$ and survival function $\bar{F}_y = 1 - F_y$. Hence, whenever a new system enters service, the time-to-failure distribution is given by the mixture distribution

$$F(t) = \sum_{y \in \mathcal{Y}} \rho_y F_y(t), \quad t \geq 0.$$

The system is inspected periodically with fixed period τ ($\tau > 0$), that is, the system is inspected at the times in the set $\{\tau, 2\tau, 3\tau, \dots\}$. At each inspection epoch, the system is observed to be in one of two states in the set $\mathcal{O} = \{0, 1\}$, where state 0 means the system is failed, and state 1 means the system is functioning. If the system is found to be in state 1, three actions are feasible: do nothing, repair, or replace. We define the set of feasible actions by $\mathcal{A} = \{0, 1, 2\}$, where 0, 1, and 2 denote do nothing, repair, and replace the system, respectively. On the other hand, if the system is found to be in state 0, only repair or replacement are permitted. Whenever a system is repaired, the same system reenters service with a virtual age of 0, and its quality type remains unaltered. If a system is replaced, a new system is randomly selected from the lot of spare systems and enters service.

Due to the indistinguishable nature of the systems, the system quality is not known with certainty. Therefore, we define a vector $\boldsymbol{b} = (b_1, b_2, \dots, b_Y)$, where b_y denotes the current belief, or probability, that $Q = y$. In other words, $\boldsymbol{b}$ is a probability distribution on the support $\mathcal{Y}$. When a new system first enters service, $\boldsymbol{b} = \boldsymbol{\rho}$, and the vector $\boldsymbol{b}$ is updated over time as the system operates. Thus, at each inspection epoch, the state of the system is described by the tuple $(x, \boldsymbol{b}) = (x, b_1, \dots, b_Y)$, where $x\tau$ is the current age of the system. For convenience, we write $x = \infty$ when the system is failed. Hence, the complete state space is given by $\mathcal{S} = \mathcal{X} \times \mathcal{Y}$, where $\mathcal{X} := \mathbb{N} \cup \{0, \infty\}$.

Denote the (random) virtual age, quality, observation, and action taken at the nth inspection time by X_n, Y_n, O_n, and A_n, respectively. At each inspection, the system is in state $S_n = (X_n, Y_n) \in \mathcal{S}$, action $A_n \in \mathcal{A}$ is taken, the state transitions to (X_{n+1}, Y_{n+1}), and the observation $O_n \in \mathcal{O}$ is received (corresponding to the new state). A cost $C(S_n, A_n)$ is incurred and the process repeats indefinitely. The cost function, C, accounts for any pertinent maintenance and downtime costs associated with taking action A_n while being in state S_n; for the moment, we leave the cost function unspecified. We seek to minimize the total expected discounted costs over an infinite time horizon. Because the state is factorable into a fully observable (age) and partially observable (quality) component, we formulate our sequential decision problem using a mixed-observability Markov decision process (MOMDP) model [16].

More formally, the MOMDP is specified by the tuple $(\mathcal{X}, \mathcal{Y}, \mathcal{A}, \mathcal{O}, P_{\mathcal{X}}, P_{\mathcal{Y}}, Z, C, \alpha)$, where $\alpha \in (0, 1)$ is the per-period discount factor and $P_{\mathcal{X}}$, $P_{\mathcal{Y}}$, and Z are functions describing the transition dynamics of the system. Specifically,

$$P_{\mathcal{X}}(x, y, a, x') = \mathbb{P}(X_{n+1} = x' \mid X_n = x, Y_n = y, A_n = a)$$

is the probability that the virtual age transitions to x' when action a is taken starting in state (x, y),

$$P_{\mathcal{Y}}(x, y, a, x', y') = \mathbb{P}(Y_{n+1} = y' \mid X_n = x, Y_n = y, A_n = a, X_{n+1} = x')$$

gives the probability that the partially observable (quality) state transitions to y' when action a is taken from state (x, y) and the fully observable state transitions to state x', and

$$Z(x', y', a, o) = \mathbb{P}(O_n = o | X_{n+1} = x', Y_{n+1} = y', A_n = a)$$

gives the probability that we receive observation o if the system transitions to state (x', y') after taking action a. Next, we define several functions that play critical roles in describing the problem's structure. Let $\bar{g}(t,y;\tau) = \mathbb{P}(T_y \geq t+\tau | T_y \geq t)$, then

$$\bar{g}(t,y;\tau) = \frac{\mathbb{P}(T_y \geq t+\tau,\, T_y \geq t)}{\mathbb{P}(T_y \geq t)} = \frac{\bar{F}_y(t+\tau)}{\bar{F}_y(t)}.$$

Because the inter-inspection time τ is fixed, we can ignore the dependence of $\bar{g}$ on τ and simply write $\bar{g}(t,y)$. Additionally, we define $g(t,y) = 1-\bar{g}(t,y)$, $\bar{G}(t,\boldsymbol{b}) = \mathbb{P}(T \geq t+\tau | T \geq t, Q \sim \boldsymbol{b})$, and $G(t,\boldsymbol{b}) = 1-\bar{G}(t,\boldsymbol{b})$, where the notation $Q \sim \boldsymbol{b}$ means that the quality type follows the probability distribution $\boldsymbol{b}$. By the law of total probability, we see that

$$\bar{G}(t,\boldsymbol{b}) = \sum_{y\in\mathcal{Y}} \mathbb{P}(T_y \geq t+\tau | T_y \geq t)\mathbb{P}(Q=y|Q\sim\boldsymbol{b}) = \sum_{y\in\mathcal{Y}} \bar{g}(t,y)b_y.$$

If no maintenance action is taken on a system in state (x,y) with $x<\infty$, it is clear that the virtual age will transition to $(x+1)\tau$ if $T_y > (x+1)\tau$; otherwise, it will transition to ∞. Therefore, for $x<\infty$,

$$P_{\mathcal{X}}(x,y,0,x') = \begin{cases} \bar{g}(x\tau,y), & x' = x+1, \\ g(x\tau,y), & x' = \infty, \\ 0, & \text{otherwise.} \end{cases}$$

However, if the system is repaired or replaced, then the age is returned to 0; hence,

$$P_{\mathcal{X}}(x,y,1,x') = P_{\mathcal{X}}(x,y,2,x') = \begin{cases} 1, & x'=0, \\ 0, & \text{otherwise.} \end{cases}$$

If the system is allowed to continue operating, or it is repaired, then its quality remains unaltered. However, if it is replaced, the new quality is again distributed according to $\boldsymbol{\rho}$. Consequently,

$$P_{\mathcal{Y}}(x,y,0,x',y') = P_{\mathcal{Y}}(x,y,1,x',y') = \begin{cases} 1, & y=y', \\ 0, & y\neq y', \end{cases}$$

and

$$P_{\mathcal{Y}}(x,y,2,x',y') = \rho_{y'}.$$

Lastly, we note that $O_n = 1$ if and only if $X_{n+1} < \infty$; hence, $Z(x',y',a,1) = \mathbb{I}(x' < \infty) = 1 - Z(x',y',a,0)$. Upon taking action a and receiving observation o, we update the belief vector using Bayes' theorem as follows:

$$\begin{aligned} b'(y'|x,\boldsymbol{b},x',a,o) &= \mathbb{P}(Q_{n+1}=y'|X_n=x, Q_n\sim\boldsymbol{b}, X_{n+1}=x', A_n=a, O_n=o) \\ &= \eta Z(x',y',a,o)\sum_{y\in\mathcal{Y}} P_{\mathcal{Y}}(x,y,a,x',y')P_{\mathcal{X}}(x,y,a,x')b_y, \end{aligned}$$

where, for any fixed $(x,\boldsymbol{b},x',a,o)$, η is a normalizing constant. Of particular interest, for $x\in\mathbb{N}$, we have

$$\bar{B}_y(x,\boldsymbol{b}) = b'(y|x,\boldsymbol{b},x+1,0,1) = \left(\bar{G}(x\tau,\boldsymbol{b})\right)^{-1}\bar{g}(x\tau,y)b_y, \tag{1}$$

and

$$B_y(x,\boldsymbol{b}) = b'(y|x,\boldsymbol{b},\infty,0,0) = \left(G(x\tau,\boldsymbol{b})\right)^{-1} g(x\tau,y)b_y,$$

for the cases when a working component is allowed to continue operating and either survives or fails, respectively. We then let $\bar{B}(x, \boldsymbol{b}) := (\bar{B}_1(x, \boldsymbol{b}), \ldots, \bar{B}_Y(x, \boldsymbol{b}))$ and $B(x, \boldsymbol{b}) := (B_1(x, \boldsymbol{b}), \ldots, B_Y(x, \boldsymbol{b}))$ as the updated belief vectors when the system survives or fails, respectively.

We assume that at each inspection, a fixed inspection cost c_I is incurred regardless of the action taken. If the system is repaired, a fixed cost c_R is incurred, but if it is replaced, a cost of $c_P > c_R$ is incurred. Lastly, if the maintenance is corrective, i.e., $x = \infty$, an additional penalty of c_F is incurred. This additional cost, c_F, reflects the fact that corrective maintenance may include additional costs such as lost production or overtime labor. The optimal total expected discounted cost, starting in belief state $(x, \boldsymbol{b})$ and denoted by $V(x, \boldsymbol{b})$, is given as a solution to the Bellman optimality equations

$$V(x, \boldsymbol{b}) = \begin{cases} \min \begin{cases} c_I + c_F + c_P + \alpha V(0, \boldsymbol{\rho}), \\ c_I + c_F + c_R + \alpha V(0, \boldsymbol{b}), \end{cases} & x = \infty, \\ \\ \min \begin{cases} c_I + c_P + \alpha V(0, \boldsymbol{\rho}), \\ c_I + c_R + \alpha V(0, \boldsymbol{b}), \\ c_I + \alpha[\bar{G}(x\tau, \boldsymbol{b}) V(x+1, \bar{B}(x+1, \boldsymbol{b})) \\ \qquad + G(x\tau, \boldsymbol{b}) V(\infty, B(x+1, \boldsymbol{b}))], \end{cases} & x < \infty. \end{cases} \tag{2}$$

The optimal action (or decision) in state $(x, \boldsymbol{b})$ is denoted by $d^*(x, \boldsymbol{b})$ so that $d^*(x, \boldsymbol{b}) = 0$ if it is optimal to do nothing, $d^*(x, \boldsymbol{b}) = 1$ if it is optimal to repair, and $d^*(x, \boldsymbol{b}) = 2$ if it is optimal to replace.

3 Structural Results

In this section, we examine the attributes of the value function and optimal policy of the MOMDP model formulated in Section 2. We begin by reviewing several stochastic orders that are used throughout our exposition. The following three definitions are adopted from the comprehensive text by Shaked and Shanthikumar [20].

Definition 1 (Usual Stochastic Order) *For two random variables X and Y, we say X is smaller than Y in the usual stochastic order (denoted $X \leq_{st} Y$) if, for all $x \in (-\infty, \infty)$,*

$$\mathbb{P}(X > x) \leq \mathbb{P}(Y > x).$$

It should be noted that, given a nondecreasing function ϕ, and assuming existence of the expectations, $X \leq_{st} Y$ implies $\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]$.

Definition 2 (Hazard Rate Order) *For two random variables X and Y, we say X is smaller than Y in the hazard rate order (denoted $X \leq_{hr} Y$) if*

$$\mathbb{P}(X > x)\mathbb{P}(Y > y) \geq \mathbb{P}(X > y)\mathbb{P}(Y > x)$$

for all real numbers x and y with $x \leq y$.

Definition 3 (Likelihood Ratio Order) *For two random variables X and Y, we say X is smaller than Y in the likelihood ratio order (denoted $X \leq_{lr} Y$) if*

$$\mathbb{P}(X \in A)\mathbb{P}(Y \in B) \geq \mathbb{P}(X \in B)\mathbb{P}(Y \in A)$$

for all measurable sets A and B such that $A \leq B$, where $A \leq B$ means that $x \in A$ and $y \in B$ imply that $x \leq y$.

The relationships between the usual stochastic, hazard rate and likelihood ratio orders are well-known and summarized as follows (see Chapter 1 of [20]): $X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y$. For convenience, if X and Y are discrete random variables with respective probability mass functions p and q, and if X and Y are ordered in a particular sense, we also say that p and q are ordered in the same sense. Furthermore, we take the converse of the statement to be true, e.g., $X \leq_{hr} Y$, if, and only if, $p \leq_{hr} q$. Next, we define the n-simplex, i.e., the space of probability mass functions on n outcomes.

Definition 4 (n-Simplex) *The standard n-simplex Δ^n is the simplex formed from the n standard unit vectors. That is,*

$$\Delta^n = \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1,\ x_i \geq 0 \text{ for all } i \right\}.$$

We also refer to Δ^n as the n-dimensional probability simplex.

Using this definition, the belief state space is $\Gamma := \mathcal{X} \times \Delta^Y$.

The results of this section require that various conditions be met. We next present these conditions and their interpretations before proceeding to the main structural results.

Condition 1 *The function $\bar{g}(t, y)$ is jointly monotone nonincreasing in t and y.*

Requiring $\bar{g}(t, y)$ to be monotone in y is necessary for meaningfully comparing the performance of systems of different qualities. Specifically, when $i < j$ at any decision epoch, a system of quality i is more likely to survive until the next decision epoch than a system of quality j. In this sense, a system from lot 1 may be regarded as being of the highest quality, and a system from lot Y of the lowest quality. This condition is equivalent to requiring that $T_y \geq_{hr} T_{y+1}$ by the equivalence of (1.B.4) and (1.B.7) in [20]. Requiring $\bar{g}(t, y)$ to be monotone in t means that for a fixed quality, as a system ages, its remaining life (stochastically) decreases. This is equivalent to saying that T_y is increasing failure rate (IFR).

Condition 2 *For each $y \in \mathcal{Y}$,*

$$\lim_{x \to \infty} \bar{g}(x\tau, y) = 0.$$

Condition 2 is a type of short-tail condition. In particular, it says that regardless of the system quality, as the system ages, it will eventually become so degraded that it will have zero probability of surviving the inter-inspection period.

Condition 3 *For each $y \neq 1$,*

$$\lim_{x \to \infty} \frac{\bar{g}(x\tau, y)}{\bar{g}(x\tau, 1)} = 0.$$

Condition 3 can be interpreted to mean that (in an asymptotic sense) lower-quality systems degrade more rapidly than the highest quality system.

We are now prepared to state our first result. It asserts that (under Condition 1), if the system survives an inter-inspection period, then the updated belief state is smaller, in the likelihood ratio, than the initial belief state. Similarly, if the system fails during an inter-inspection period, then the updated belief state is larger in the likelihood ratio than the initial belief state.

Proposition 1 *Under Condition 1, for $x \in \mathcal{X}$ and $\boldsymbol{b} \in \Delta^Y$*

$$\bar{B}(x, \boldsymbol{b}) \leq_{lr} \boldsymbol{b} \leq_{lr} B(x, \boldsymbol{b}).$$

Proof. First, we show that $\bar{B}(x, \boldsymbol{b}) \leq_{lr} \boldsymbol{b}$. To this end, we must show $b_y/\bar{B}_y(x, \boldsymbol{b})$ is nondecreasing in $y \in \mathcal{Y}$. By equation (1), we have

$$\frac{b_y}{\bar{B}_y(x, \boldsymbol{b})} = \frac{\bar{G}(x\tau, \boldsymbol{b})b_y}{\bar{g}(x\tau, y)b_y} = \frac{\bar{G}(x\tau, \boldsymbol{b})}{\bar{g}(x\tau, y)}.$$

The proof is completed by noting that Condition 1 implies $\bar{g}(t, y)$ is monotone nonincreasing in $y \in \mathcal{Y}$. The proof for $\bar{B}(x, \boldsymbol{b}) \leq_{lr} \boldsymbol{b}$ follows similarly by noting that $g(t, y) = 1 - \bar{g}(t, y)$ is monotone nondecreasing in $y \in \mathcal{Y}$. ∎

We next show that, for any fixed belief state, the probability that the system survives the inter-inspection period decreases as the virtual age increases; consequently, the probability of failure increases.

Proposition 2 *Under Condition 1, for $\boldsymbol{b} \in \Delta^Y$ and $x \in \mathcal{X} \setminus \{\infty\}$,*

$$\bar{G}(x\tau, \boldsymbol{b}) \geq \bar{G}((x+1)\tau, \boldsymbol{b}) \quad \text{and} \quad G(x\tau, \boldsymbol{b}) \leq G((x+1)\tau, \boldsymbol{b}).$$

Proof. We will first show that $\bar{G}(x\tau, \boldsymbol{b}) \geq \bar{G}((x+1)\tau, \boldsymbol{b})$, noting that $G(x\tau, \boldsymbol{b}) \leq G((x+1)\tau, \boldsymbol{b})$ follows since $G(t, \boldsymbol{b}) = 1 - \bar{G}(t, \boldsymbol{b})$. By Condition 1, for each $x \in \mathcal{X} \setminus \{\infty\}$,

$$h_1(y) := \bar{g}(x\tau, y) \geq \bar{g}((x+1)\tau, y) := h_2(y)$$

for all $y \in \mathcal{Y}$. Because the functions h_1 and h_2 are ordered for all y,

$$h_1(Q) \geq_{st} h_2(Q). \tag{3}$$

Consequently, by taking expectations on across both sides of inequality (3),

$$\bar{G}(x\tau, \boldsymbol{b}) = \mathbb{E}(h_1(Q)) \geq \mathbb{E}(h_2(Q)) = \bar{G}((x+1)\tau, \boldsymbol{b}),$$

where $Q \sim \boldsymbol{b}$. ∎

The next result asserts that if two states are ordered such that the belief state is ordered in the sense of the usual stochastic order, and the virtual ages are ordered, then the probabilities of system failure are ordered in the same direction. That is, systems with older virtual ages and stochastically larger belief states are more likely to fail.

Proposition 3 *Under Conditions 1 and 2, if $\boldsymbol{b}_1 \leq_{st} \boldsymbol{b}_2$ and $x_1 \leq x_2$ then*

$$\bar{G}(x_1\tau, \boldsymbol{b}_1) \geq \bar{G}(x_2\tau, \boldsymbol{b}_2) \quad \text{and} \quad G(x_1\tau, \boldsymbol{b}_1) \leq G(x_2\tau, \boldsymbol{b}_2).$$

Proof. We again proceed by first showing $\bar{G}(x_1\tau, \boldsymbol{b}_1) \geq \bar{G}(x_2\tau, \boldsymbol{b}_2)$, noting this implies $G(x_1\tau, \boldsymbol{b}_1) \leq G(x_2\tau, \boldsymbol{b}_2)$. By Condition 1, we have that $\bar{g}(x_1\tau, y)$ is nonincreasing in $y \in \mathcal{Y}$. Thus, if $Q_i \sim \boldsymbol{b}_i$, $i = 1, 2$, then

$$\bar{G}(x_1\tau, \boldsymbol{b}_1) = \mathbb{E}(\bar{g}(x_1\tau, Q_1)) \geq \mathbb{E}(\bar{g}(x_1\tau, Q_2)) = \bar{G}(x_2\tau, \boldsymbol{b}_2)$$

where the inequality holds by $\boldsymbol{b}_1 \leq_{st} \boldsymbol{b}_2$. The result follows by repeated application of Proposition 2. ∎

Next, assume that we have two distributions over the quality of the system in operation that are ordered in the likelihood ratio sense. Under the same observation, i.e., the system is found to be functioning or failed at a particular time, the distributions remain ordered after being updated. This is formalized in Proposition 4.

Proposition 4 *Under Condition 1, if $\boldsymbol{b}_1 \leq_{lr} \boldsymbol{b}_2$ and $x \in \mathcal{X}$ then*

$$\bar{B}(x, \boldsymbol{b}_1) \leq_{lr} \bar{B}(x, \boldsymbol{b}_2) \quad \text{and} \quad B(x, \boldsymbol{b}_1) \leq_{lr} B(x, \boldsymbol{b}_2).$$

Proof. Again, we proceed by only showing that $\bar{B}(x, \boldsymbol{b}_1) \leq_{lr} \bar{B}(x, \boldsymbol{b}_2)$. For $i = 1, 2$, let $\boldsymbol{b}_y^i$ be the yth component of $\boldsymbol{b}_i$. Then,

$$\frac{\bar{B}_y(x, \boldsymbol{b}_2)}{\bar{B}_y(x, \boldsymbol{b}_1)} = \frac{[\bar{G}(x\tau, \boldsymbol{b}_2)]^{-1}}{[\bar{G}(x\tau, \boldsymbol{b}_1)]^{-1}} \frac{\bar{g}(x\tau, y)}{\bar{g}(x\tau, y)} \frac{\boldsymbol{b}_y^2}{\boldsymbol{b}_y^1} = \frac{\bar{G}(x\tau, \boldsymbol{b}_1)}{\bar{G}(x\tau, \boldsymbol{b}_2)} \frac{\boldsymbol{b}_y^2}{\boldsymbol{b}_y^1},$$

but $\boldsymbol{b}_y^2/\boldsymbol{b}_y^1$ is increasing in y by assumption; hence, $\bar{B}_y(x, \boldsymbol{b}_2)/\bar{B}_y(x, \boldsymbol{b}_1)$ is increasing. ∎

We are now prepared to state our first main result. Theorem 1 asserts that, under Condition 1, for a fixed virtual age, the value function is monotone in the belief state. Additionally, for any fixed belief state, the value function is largest when the system is in the failed state.

Theorem 1 *Under Condition 1, if $x \in \mathcal{X}$ and $\boldsymbol{b}_1 \leq_{lr} \boldsymbol{b}_2$, then*

$$V(x, \boldsymbol{b}_1) \leq V(x, \boldsymbol{b}_2)$$

and if $\boldsymbol{b} \in \Delta^Y$ then

$$c_F + V(x, \boldsymbol{b}) \leq V(\infty, \boldsymbol{b}).$$

Proof. We proceed by induction on the iterates of the value iteration algorithm. Let v^k be the approximation of the value function V at the kth iteration and assume $v^0(x, \boldsymbol{b}) = 0$ for all $(x, \boldsymbol{b}) \in \mathcal{X} \times \Delta^Y$. Then,

$$v^1(x, \boldsymbol{b}) = \begin{cases} c_I + c_F + c_R, & x = \infty, \\ c_I, & x < \infty; \end{cases}$$

hence, $v^1(x, \boldsymbol{b})$ is constant in $\boldsymbol{b}$ and $v^1(x, \boldsymbol{b}) + c_F < v^1(\infty, \boldsymbol{b})$. Now, assume that $v^k(x, \boldsymbol{b})$ satisfies the induction hypothesis, i.e., if $x \in \mathcal{X}$ and $\boldsymbol{b}_1 \leq_{lr} \boldsymbol{b}_2$ then $v^k(x, \boldsymbol{b}_1) \leq v^k(x, \boldsymbol{b}_2)$, and if $b \in \Delta^Y$ then

$c_F + v^k(x,\boldsymbol{b}) \le v^k(x,\boldsymbol{b})$. For notational convenience, we let $C_P^k = c_I + c_P + \alpha v^k(0,\boldsymbol{\rho})$, $C_R^k(\boldsymbol{b}) = c_I + c_R + \alpha v^k(0,\boldsymbol{b})$, and $C_{DN}^k(x,\boldsymbol{b}) = c_I + \alpha(\bar{G}(x,\boldsymbol{b})v^k(x+1,\bar{B}(x+1,\boldsymbol{b})) + G(x,\boldsymbol{b})v^k(\infty, B(x+1,\boldsymbol{b}))$. Then, when $x=\infty$, we have

$$v^{k+1}(\infty,\boldsymbol{b}) = \min\{c_F + C_P^k,\, c_F + C_R^k(\boldsymbol{b})\}.$$

By the induction hypothesis, $C_R^k(\boldsymbol{b})$ is monotone nondecreasing in $\boldsymbol{b}$, and since C_P^k is constant in $\boldsymbol{b}$, it is clear that $v^{k+1}(\infty,\boldsymbol{b})$ is also monotone nondecreasing in $\boldsymbol{b}$.

Next, for $x<\infty$,

$$v^{k+1}(x,\boldsymbol{b}) = \min\{C_P^k,\, C_R^k(\boldsymbol{b}),\, C_{DN}^k(x,\boldsymbol{b})\} \le \min\{C_P^k,\, C_R^k(\boldsymbol{b})\};$$

hence, $v^{k+1}(x,\boldsymbol{b}) + c_F \le c_F + \min\{C_P^k, C_R^k(\boldsymbol{b})\} = v^{k+1}(\infty,\boldsymbol{b})$. To complete the proof, we need to show that $C_{DN}^k(x,\boldsymbol{b})$ is monotone nondecreasing in $\boldsymbol{b}$. Let, $\boldsymbol{b}_1 \le_{lr} \boldsymbol{b}_2$, and, for $i=1,2$, let D_i be a random variable such that

$$\mathbb{P}(D_i = q) = \begin{cases} \bar{G}(x,\boldsymbol{b}_i), & q=0, \\ G(x,\boldsymbol{b}_i), & q=1, \end{cases}$$

and let $h_i^k(q|x)$ be a function such that

$$h_i^k(q|x) = \begin{cases} v^k(x+1,\bar{B}(x+1,\boldsymbol{b}_i)), & q=0, \\ v^k(\infty, B(x+1,\boldsymbol{b}_i)), & q=1. \end{cases}$$

Because $\boldsymbol{b}_1 \le_{lr} \boldsymbol{b}_2$ implies $\boldsymbol{b}_1 \le_{st} \boldsymbol{b}_2$, Proposition 2 shows that $G(x,\boldsymbol{b}_1) \le G(x,\boldsymbol{b}_2)$; consequently, $D_1 \le_{st} D_2$. Additionally, by Proposition 4, $\bar{B}(x+1,\boldsymbol{b}_1) \le_{lr} \bar{B}(x+1,\boldsymbol{b}_2)$ and $B(x+1,\boldsymbol{b}_1) \le_{lr} B(x+1,\boldsymbol{b}_2)$; therefore, by the induction hypothesis, $h_1^k(q|x) \le h_2^k(q|x)$ for $q=0,1$. We can see then that

$$C_{DN}^k(x,\boldsymbol{b}_1) = c_I + \alpha\,\mathbb{E}(h_1^k(D_1|x)) \le c_I + \alpha\,\mathbb{E}(h_2^k(D_1|x)).$$

Next, we note

$$\begin{aligned} h_2^k(0|x) = v^k(x+1,\bar{B}(x+1,\boldsymbol{b}_2)) \le V^k(x+1,\boldsymbol{b}_2) &\le v^k(x+1,B(x+1,\boldsymbol{b}_2)) \\ &\le v^k(\infty, B(x+1,\boldsymbol{b}_2)) \\ &= h_2^k(1|x), \end{aligned}$$

where the first two inequalities follow from the induction hypothesis and Proposition 1, and the last inequality follows from the induction hypothesis. Thus, $h_2^k(q|x)$ is nondecreasing in q, so by $D_1 \le_{st} D_2$

$$C_{DN}^k(x,\boldsymbol{b}_1) \le c_I + \alpha\,\mathbb{E}(h_2^k(D_1|x)) \le c_I + \alpha\,\mathbb{E}(h_2^k(D_2|x)) = C_{DN}^k(x,\boldsymbol{b}_2),$$

and the proof is complete. ■

Our next result (Theorem 2) establishes that the belief state space can be partitioned into two regions: a region where either doing nothing or repair is optimal, and a region where either doing nothing or replacement is optimal. Moreover, these regions are formed by partitioning Δ^Y and are related to the likelihood ratio ordering.

Theorem 2 *Under Condition 1, if $d^*(x, \boldsymbol{b}) = 2$, then $d^*(x', \boldsymbol{b}') \in \{0, 2\}$ for all $x' \in \mathbb{N}$ and $\boldsymbol{b}' \geq_{lr} \boldsymbol{b}$. Additionally, if $d^*(x, \boldsymbol{b}) = 1$, then $d^*(x', \boldsymbol{b}') \in \{0, 1\}$ for all $x' \in \mathbb{N}$ and $\boldsymbol{b}' \leq_{lr} \boldsymbol{b}$.*

Proof. If $d^*(x, \boldsymbol{b}) = 2$, then $V(x, \boldsymbol{b}) = c_I + c_P + \alpha V(0, \boldsymbol{\rho})$; hence,

$$\begin{aligned} c_I + c_P + \alpha V(0, \boldsymbol{\rho}) &< c_I + c_R + \alpha V(0, \boldsymbol{b}) \qquad (4) \\ &\leq c_I + c_R + \alpha V(0, \boldsymbol{b}'), \qquad (5) \end{aligned}$$

where the first inequality follows directly from (2), and the second inequality follows from Theorem 1. Therefore, repair is not optimal in state $(x', \boldsymbol{b}')$. The proof of the second statement follows by noting that $d^*(x, \boldsymbol{b}) = 1$ and $\boldsymbol{b}' \leq_{lr} \boldsymbol{b}$ results in a reversal of inequalities (4) and (5). ■

To further understand the structure of the optimal value function and policy, we consider their behavior when the quality is known with certainty. Let the set of vectors $\{\boldsymbol{e}_y : 1 \leq y \leq Y\}$ be the standard basis for Y-dimensional Euclidean space, where $\boldsymbol{e}_y$ denotes the vector with a one in the yth coordinate and zeros elsewhere. Our next result states that, when the quality is known with certainty, the value function is monotone nondecreasing in the virtual age.

Proposition 5 *Under Condition 1, for each $y \in \mathcal{Y}$, and $x \in \mathcal{X} \setminus \{\infty\}$,*

$$V(x, \boldsymbol{e}_y) \leq V(x+1, \boldsymbol{e}_y).$$

Proof. Again, by induction, if $v^0(x, \boldsymbol{b}) = 0$ for all $(x, \boldsymbol{b}) \in \mathcal{X} \times \Delta^Y$, then again

$$v^1(x, \boldsymbol{b}) = \begin{cases} c_I + c_R + c_F, & x = \infty, \\ c_I, & x < \infty; \end{cases}$$

hence, the base case holds. Additionally, for each $x \in \mathcal{X} \setminus \{\infty\}$ and $y \in \mathcal{Y}$, we note that $v^1(x, \boldsymbol{e}_y) \leq v^1(\infty, \boldsymbol{e}_y)$. Now, assume that for each x, y that $v^k(x, \boldsymbol{e}_y) \leq v^k(x+1, \boldsymbol{e}_y) \leq v^k(\infty, \boldsymbol{e}_y)$, which also holds in the base case, then by the proof of Theorem 1 we know this property is conserved under each iteration of the value iteration algorithm and holds for $k+1$. Then, noting that $\bar{G}(x, \boldsymbol{e}_y) = \bar{g}(x, y)$, $G(x, \boldsymbol{e}_y) = g(x, y)$, and $\bar{B}(x, \boldsymbol{e}_y) = B(x, \boldsymbol{e}_y) = \boldsymbol{e}_y$, we see

$$\begin{aligned} v^{k+1}(x, \boldsymbol{e}_y) &= \min \begin{cases} c_I + \alpha v^k(0, \boldsymbol{\rho}) \\ c_I + \alpha v^k(0, \boldsymbol{e}_y) \\ c_I + \alpha \left(\bar{g}(x, y) v^k(x+1, \boldsymbol{e}_y) + g(x, y) v^k(\infty, \boldsymbol{e}_y) \right) \end{cases} \\ &\leq \min \begin{cases} c_I + \alpha v^k(0, \boldsymbol{\rho}) \\ c_I + \alpha v^k(0, \boldsymbol{e}_y) \\ c_I + \alpha \left(\bar{g}(x, y) v^k(x+2, \boldsymbol{e}_y) + g(x, y) v^k(\infty, \boldsymbol{e}_y) \right) \end{cases} \\ &\leq \min \begin{cases} c_I + \alpha v^k(0, \boldsymbol{\rho}) \\ c_I + \alpha v^k(0, \boldsymbol{e}_y) \\ c_I + \alpha \left(\bar{g}(x+1, y) v^k(x+2, \boldsymbol{e}_y) + g(x+1, y) v^k(\infty, \boldsymbol{e}_y) \right) \end{cases} \\ &= v^{k+1}(x+1, \boldsymbol{e}_y), \end{aligned}$$

where the first inequality follows directly from the induction hypothesis, and the second inequality follows by noting that $v^k(x+2, \boldsymbol{e}_y) \leq v^k(\infty, \boldsymbol{e}_y)$ and $g(x, y) \leq g(x+1, y)$ (using a stochastic ordering argument similar to that in the proof of Theorem 1). ■

Additionally, for each $\boldsymbol{e}_y$, there exists a virtual age above which it is optimal to preventively maintain and below which it is optimal to do nothing. This result is formalized in Proposition 6.

Proposition 6 *If $d^*(x, \boldsymbol{e}_y) > 0$, then*

$$V(x+1, \boldsymbol{e}_y) = V(x, \boldsymbol{e}_y) = V(\infty, \boldsymbol{e}_y) - c_F,$$

and

$$d^*(x+1, \boldsymbol{e}_y) = d^*(x, \boldsymbol{e}_y).$$

Proof. If $d^*(x, \boldsymbol{e}_y) > 0$, then

$$\begin{aligned} V(x, \boldsymbol{e}_y) &= \min\{c_I + c_P + \alpha V(0, \boldsymbol{\rho}), c_I + c_R + \alpha V(0, \boldsymbol{e}_y)\} \\ &\leq c_I + \alpha \left(\bar{g}(x, y) V(x+1, \boldsymbol{e}_y) + g(x, y) V(\infty, \boldsymbol{e}_y)\right) \quad (6) \\ &\leq c_I + \alpha \left(\bar{g}(x, y) V(x+2, \boldsymbol{e}_y) + g(x, y) V(\infty, \boldsymbol{e}_y)\right) \quad (7) \\ &\leq c_I + \alpha \left(\bar{g}(x+1, y) V(x+2, \boldsymbol{e}_y) + g(x+1, y) V(\infty, \boldsymbol{e}_y)\right), \quad (8) \end{aligned}$$

where inequality (6) follows by supposition that the optimal action is preventive maintenance, (7) follows by Proposition 5, and (8) follows by Condition 1. The result then follows immediately by noting that

$$\begin{aligned} V(x+1, \boldsymbol{e}_y) &= \min \begin{cases} \min\{c_I + c_P + \alpha V(0, \boldsymbol{\rho}), c_I + c_R + \alpha V(0, \boldsymbol{e}_y)\} \\ c_I + \alpha \left(\bar{g}(x+1, y) V(x+2, \boldsymbol{e}_y) + g(x+1, y) V(\infty, \boldsymbol{e}_y)\right) \end{cases} \\ &= \min\{c_I + c_P + \alpha V(0, \boldsymbol{\rho}), c_I + c_R + \alpha V(\infty, \boldsymbol{e}_y)\} \\ &= V(x, \boldsymbol{e}_y). \end{aligned}$$

■

Proposition 7, further characterizes the optimal policy and value function when the quality is known with certainty.

Proposition 7 *Suppose that Conditions 1 and 2 hold. For each $y \in \mathcal{Y}$,*

1. *if $d^*(x, \boldsymbol{e}_y) \in \{0, 2\}$, then $d^*(x, \boldsymbol{e}_y) = 0$ for all $x < x_y^*$ and $d^*(x, \boldsymbol{e}_y) = 2$ for all $x \geq x_y^*$, where*

$$x_y^* = \operatorname{argmin}\left\{x \in \mathbb{N} : \bar{g}(x\tau, y) < \left(\frac{1-\alpha}{\alpha c_F}\right)\left(\frac{c_I + c_F}{1-\alpha} - (c_I + c_F + c_P) - V(0, \boldsymbol{\rho})\right)\right\},$$

2. *if $d^*(x, \boldsymbol{e}_y) \in \{0, 1\}$, then $d^*(x, \boldsymbol{e}_y) = 0$ for all $x < x_y^*$ and $d^*(x, \boldsymbol{e}_y) = 1$ for all $x \geq x_y^*$, where*

$$x_y^* = \operatorname{argmin}\left\{x \in \mathbb{N} : \bar{g}(x\tau, y) < \left(\frac{1-\alpha}{\alpha c_F}\right)\left(\frac{c_I + c_F}{1-\alpha} - (c_I + c_F + c_R) - V(0, \boldsymbol{e}_y)\right)\right\},$$

3. if $d^*(x, \boldsymbol{e}_y) \in \{0,1\}$, then $V(x, \boldsymbol{e}_y) = V(x)$ where $V(x)$ satisfies the following one-dimensional Bellman equations:

$$V(x) = \min \begin{cases} c_I + c_P + \alpha V(0), \\ c_I + \alpha \left[\bar{g}(x\tau, y) V(x+1) + g(x\tau, y) \left(c_I + c_F + c_R + \alpha V(0) \right) \right]. \end{cases}$$

Proof. We first prove Part 3. By the assumption that $d^*(x, \boldsymbol{e}_y) \in \{0,1\}$, and (2), we have

$$V(x, \boldsymbol{e}_y) = \min \begin{cases} c_I + c_R + \alpha V(0, \boldsymbol{e}_y), \\ c_I + \alpha \left[\bar{G}(x\tau, \boldsymbol{e}_y) V(x+1, \bar{B}(x+1, \boldsymbol{e}_y)) + G(x\tau, \boldsymbol{e}_y) V(\infty, B(x+1, \boldsymbol{e}_y)) \right]. \end{cases} \tag{9}$$

It can be seen that $\boldsymbol{e}_y = \bar{B}(x, \boldsymbol{e}_y) = B(x, \boldsymbol{e}_y)$, $\bar{G}(x\tau, \boldsymbol{e}_y) = \bar{g}(x\tau, y)$ and $G(x\tau, \boldsymbol{e}_y) = g(x\tau, y)$. Hence, $V(\infty, B(x+1, \boldsymbol{e}_y)) = V(\infty, \boldsymbol{e}_y) = c_I + c_F + c_R + \alpha V(0, \boldsymbol{e}_y)$ and (9) can be rewritten as

$$V(x, \boldsymbol{e}_y) = \min \begin{cases} c_I + c_R + \alpha V(0, \boldsymbol{e}_y), \\ c_I + \alpha \left[\bar{g}(x\tau, y) V(x+1, \boldsymbol{e}_y) + g(x\tau, y)(c_I + c_F + c_R + \alpha V(0, \boldsymbol{e}_y)) \right], \end{cases} \tag{10}$$

completing the proof of Part 3. For Part 2, we see from (10) that $d^*(x, \boldsymbol{e}_y) = 1$ if, and only if,

$$c_I + c_R + \alpha V(0, \boldsymbol{e}_y) < c_I + \alpha \left[\bar{g}(x\tau, y) V(x+1, \boldsymbol{e}_y) + g(x\tau, y)(c_I + c_F + c_R + \alpha V(0, \boldsymbol{e}_y)) \right]. \tag{11}$$

We note that if $d^*(x, \boldsymbol{e}_y) = 1$, then by Proposition 6 that $d^*(x+1, \boldsymbol{e}_y) = 1$, and, consequently, $V(x, \boldsymbol{e}_y) = V(x+1, \boldsymbol{e}_y) = c_I + c_R + \alpha V(0, \boldsymbol{e}_y)$. Hence, inequality (11) is equivalent to

$$c_I + c_R + \alpha V(0, \boldsymbol{e}_y) < c_I + \alpha \left[\bar{g}(x\tau, y)(c_I + c_R + \alpha V(0, \boldsymbol{e}_y)) + g(x\tau, y)(c_I + c_F + c_R + \alpha V(0, \boldsymbol{e}_y)) \right]. \tag{12}$$

By noting that $g(x\tau, y) + \bar{g}(x\tau, y) = 1$, we rearrange the inequality to see that $d^*(x, \boldsymbol{e}_y) = 1$ if, and only if,

$$\bar{g}(x\tau, y) < \left(\frac{1-\alpha}{\alpha c_F} \right) \left(\frac{c_I + c_F}{1-\alpha} - (c_I + c_F + c_R) - V(0, \boldsymbol{e}_y) \right). \tag{13}$$

The proof is completed by noting that $\bar{g}(x\tau, y)$ is monotone nonincreasing in x, and converges to 0 (by Conditions 1 and 2, respectively). The proof of Part 1 is similar to that of Part 2. ∎

Proposition 7, Part 3 is useful for computing the optimal policy at extreme points of Δ^Y where repair is optimal. Additionally, as shown in Corollary 1, Proposition 7, Part 1 is useful for bounding the age replacement threshold of the extreme points of Δ^Y for which repair is not optimal.

Corollary 1 *Under Conditions 1-2, for each $y \in \mathcal{Y}$, if $d^*(x, \boldsymbol{e}_y) \in \{0,2\}$, and if there exist real numbers $\underline{V}$ and $\bar{V}$ such that $\underline{V} \le V(0, \boldsymbol{\rho}) \le \bar{V}$, then $\underline{x}_y \le x^*_y \le \bar{x}_y$, where*

$$\underline{x}_y = \operatorname{argmin} \left\{ x \in \mathbb{N} : \bar{g}(x\tau, y) < \left(\frac{1-\alpha}{\alpha c_F} \right) \left(\frac{c_I + c_F}{1-\alpha} - (c_I + c_F + c_P) - \underline{V} \right) \right\}, \text{ and}$$

$$\bar{x}_y = \operatorname{argmin} \left\{ x \in \mathbb{N} : \bar{g}(x\tau, y) < \left(\frac{1-\alpha}{\alpha c_F} \right) \left(\frac{c_I + c_F}{1-\alpha} - (c_I + c_F + c_P) - \bar{V} \right) \right\}.$$

There are many ways to bound $V(0,\boldsymbol{\rho})$, but some simple bounds are given by $V(0,\boldsymbol{e}_1) \leq V(0,\boldsymbol{\rho}) \leq \underline{V}(0,\boldsymbol{\rho})$ where $V(0,\boldsymbol{e}_1)$ can be computed as described in Proposition 7, Part 3 and $\underline{V}(0,\boldsymbol{\rho})$ is determined by solving the following one-dimensional Bellman equations:

$$\underline{V}(x,\boldsymbol{\rho}) = \min \begin{cases} c_I + c_P + \alpha \underline{V}(0,\boldsymbol{\rho}), \\ c_I + \alpha\left(\bar{G}(x\tau,\boldsymbol{\rho})\underline{V}(x+1,\boldsymbol{\rho}) + \bar{G}(x\tau,\boldsymbol{\rho})(c_I + c_F + c_P + \alpha\underline{V}(0,\boldsymbol{\rho}))\right). \end{cases} \tag{14}$$

The solution to equation (14) gives the total expected discounted cost of an optimal replacement-only policy when the belief about the active system's quality is never updated.

In order to prove our final main result concerning the structure of the optimal policy, we need several useful lemmas, whose proofs are provided in the Appendix. The first lemma establishes limits for the Bayesian update functions B and $\bar{B}$.

Lemma 1 *For all $\boldsymbol{b} \in \Delta^Y$, under Condition 2,*

$$\lim_{x\to\infty} B(x,\boldsymbol{b}) = \boldsymbol{b},$$

and, under Condition 3,

$$\lim_{x\to\infty} \bar{B}(x,\boldsymbol{b}) = \boldsymbol{e}_1.$$

Lemma 2 asserts that, regardless of the time-to-failure distribution, $\bar{G}$, G, $\bar{B}$, and B are continuous in the belief about the system quality.

Lemma 2 *For each fixed $x \in \mathcal{X} \setminus \{\infty\}$, the functions $\bar{G}(x,\boldsymbol{b})$, $G(x,\boldsymbol{b})$, $\bar{B}(x,\boldsymbol{b})$, and $B(x,\boldsymbol{b})$ are continuous in $\boldsymbol{b} \in \Delta^Y$.*

Our next result states that, for each fixed virtual age, the optimal value function V is continuous in the belief about the system quality.

Lemma 3 *Under Conditions 2-3, for each fixed $x \in \mathcal{X} \setminus \{\infty\}$, $V(x,\boldsymbol{b})$ is continuous in $\boldsymbol{b} \in \Delta^Y$.*

The following two results enable us to characterize the asymptotic behavior of the value function as the virtual age increases. Specifically, Lemma 4 quantifies the value of doing nothing as the virtual age goes to infinity, while Corollary 2 asserts that the value function is convergent as $x \to \infty$.

Lemma 4 *Under Conditions 2-3,*

$$\lim_{x\to\infty} V_{DN}(x,\boldsymbol{b}) = c_I + \alpha V(\infty,\boldsymbol{b}),$$

where $V_{DN}(x,\boldsymbol{b}) = c_I + \alpha\left(\bar{G}(x,\boldsymbol{b})V(x+1,\bar{B}(x+1,\boldsymbol{b})) + G(x,\boldsymbol{b})V(\infty,B(x+1,\boldsymbol{b}))\right)$.

A natural consequence of Lemma 4 is that, for each fixed system belief, the value function converges as $x \to \infty$. We state Corollary 2 without proof.

Corollary 2 *Under Conditions 2-3,*

$$\lim_{x\to\infty} V(x,\boldsymbol{b}) = c_I + \min\{c_P + \alpha V(0,\boldsymbol{\rho}), c_R + \alpha V(0,\boldsymbol{b}), \alpha V(\infty,\boldsymbol{b})\}.$$

Finally, Theorem 3 asserts that, for each fixed system belief $\boldsymbol{b} \in \Delta^Y$, under the appropriate conditions, there exists a threshold in the virtual age beyond which preventive maintenance is optimal. Moreover, this threshold is guaranteed to be finite.

Theorem 3 *Under Conditions 1-3, if $c_R < \alpha c_F$, then for each $\boldsymbol{b} \in \Delta^Y$, there exists an $x(\boldsymbol{b}) < \infty$ such that $d^*(x, \boldsymbol{b}) = d^*(x(\boldsymbol{b}), \boldsymbol{b}) > 0$ for all $x > x(\boldsymbol{b})$.*

Proof. By Theorem 1,

$$\alpha V(\infty, \boldsymbol{b}) \geq \alpha(c_F + V(x, \boldsymbol{b})),$$

for all $x \in \mathcal{X}$. Therefore,

$$\begin{aligned} \alpha V(\infty, \boldsymbol{b}) \geq \alpha(c_F + V(0, \boldsymbol{b})) &= \alpha c_F + \alpha V(0, \boldsymbol{b}) && (15)\\ &> c_R + \alpha V(0, \boldsymbol{b}) && (16)\\ &\geq \min\{c_P + \alpha V(0, \boldsymbol{\rho}), c_R + \alpha V(0, \boldsymbol{b})\}, && (17) \end{aligned}$$

where inequality (16) follows from the assumption that $c_R < \alpha c_F$, and inequality (17) by definition of the minimum function. For all $\delta, \epsilon > 0$, there exist two finite, possibly distinct, integers $x_{DN}(\delta, \boldsymbol{b})$ and $x(\epsilon, \boldsymbol{b})$ such that for all $x > x_{DN}(\delta, \boldsymbol{b})$,

$$|V_{DN}(x, \boldsymbol{b}) - (c_I + \alpha V(\infty, \boldsymbol{b}))| < \delta, \tag{18}$$

and for all $x > x(\epsilon, \boldsymbol{b})$

$$|V(x, \boldsymbol{b}) - (c_I + \min\{c_P + \alpha V(0, \boldsymbol{\rho}), c_R + \alpha V(0, \boldsymbol{b})\})| < \epsilon, \tag{19}$$

where (18) follows from Lemma 4, and (19) follows from Corollary 2 and the strict inequality (16).

Define

$$d(\boldsymbol{b}) = c_I + \alpha V(\infty, \boldsymbol{b}) - c_I + \min\{c_P + \alpha V(0, \boldsymbol{\rho}), c_R + \alpha V(0, \boldsymbol{b})\},$$

then for any $x > \max\{x_{DN}(d(\boldsymbol{b})/2, \boldsymbol{b}), x(d(\boldsymbol{b})/2, \boldsymbol{b})\}$, we have

$$\begin{aligned} V_{DN}(x, \boldsymbol{b}) - V(x, \boldsymbol{b}) &= V_{DN}(x, \boldsymbol{b}) - V(x\boldsymbol{b}) + d(\boldsymbol{b}) - d(\boldsymbol{b}) && (20)\\ &= V_{DN}(x, \boldsymbol{b}) - V(x\boldsymbol{b}) + d(\boldsymbol{b}) - [c_I + \alpha V(\infty, \boldsymbol{b}) && (21)\\ &\qquad\qquad - c_I + \min\{c_P + \alpha V(0, \boldsymbol{\rho}), c_R + \alpha V(0, \boldsymbol{b})\}] \\ &= (V_{DN}(x, \boldsymbol{b}) - c_I + \alpha V(\infty, \boldsymbol{b}) && (22)\\ &\quad + [c_I + \min\{c_P + \alpha V(0, \boldsymbol{\rho}), c_R + \alpha V(0, \boldsymbol{b})\} - V(x, \boldsymbol{b})] + d(\boldsymbol{b}) \\ &> -\frac{d(\boldsymbol{b})}{2} - \frac{d(\boldsymbol{b})}{2} + d(\boldsymbol{b}) && (23)\\ &= 0, && (24) \end{aligned}$$

where inequality (23) follows by (18), (19), and the fact that (15)-(17) imply $d > 0$. Thus, for all $x > \max\{x_{DN}(d/2, \boldsymbol{b}), x(d/2, \boldsymbol{b})\}$, doing nothing is strictly suboptimal. The result follows by defining $x(\boldsymbol{b}) = \max\{x_{DN}(d(\boldsymbol{b})/2, \boldsymbol{b}), x(d(\boldsymbol{b})/2, \boldsymbol{b})\}$. ∎

We are unable to confirm that the value function is jointly monotone nondecreasing; however, if such monotonicity holds, we can establish stronger structural properties.

Theorem 4 *If $V(x, \boldsymbol{b})$ is jointly monotone nondecreasing, then*

1. *for each $\boldsymbol{b} \in \Delta^Y$, $x \leq x'$ implies $d^*(x, \boldsymbol{b}) \leq d^*(x', \boldsymbol{b})$; and*

2. *if $d^*(x_1, \boldsymbol{b}_1) = 2$, then $d^*(x_2, \boldsymbol{b}_2) = 2$ for all $x_2 \geq x_1$ and $\boldsymbol{b}_2 \geq_{lr} \boldsymbol{b}_1$.*

Proof. Part 1: By contradiction, assume that for some $(x, \boldsymbol{b}) \in \mathcal{X} \times \Delta^Y$ and $x' > x$ that $d^*(x, \boldsymbol{b}) > d^*(x', \boldsymbol{b})$. It follows by Theorem 2 that $V(x, \boldsymbol{b}) = \min\{c_I + c_R + \alpha V(0, \boldsymbol{b}), c_I + c_R + \alpha V(0, \boldsymbol{\rho})\}$ and $V(x', \boldsymbol{b}) = \min\{c_I + c_R + \alpha V(0, \boldsymbol{b}), c_I + c_R + \alpha V(0, \boldsymbol{\rho}), V_D N x', \boldsymbol{b}\} \leq \min\{c_I + c_R + \alpha V(0, \boldsymbol{b}), c_I + c_R + \alpha V(0, \boldsymbol{\rho})\} = V(x, \boldsymbol{b})$. Therefore, $V(x, \boldsymbol{b}) = V(x', \boldsymbol{b})$ and $d^*(x, \boldsymbol{b}) = d^*(x', \boldsymbol{b})$, which is a contradiction.

Part 2: If $d^*(x_1, \boldsymbol{b}_1) = 2$, then by equation (2) we have that $V(x_1, \boldsymbol{b}_1) = c_I + c_P + \alpha V(0, \boldsymbol{\rho})$, and by joint monotonicity it is seen that $V(x_1, \boldsymbol{b}_1) \leq V(x_2, \boldsymbol{b}_2)$. Lastly, we note that $V(x_2, \boldsymbol{b}_2) \leq c_I + c_P + \alpha V(0, \boldsymbol{\rho})$. Combining these facts yields

$$c_I + c_P + \alpha V(0, \boldsymbol{\rho}) \leq V(x_2, \boldsymbol{b}_2) \leq c_I + c_P + \alpha V(0, \boldsymbol{\rho}). \tag{25}$$

Therefore, equality holds throughout (25), thus completing the proof. ∎

Theorem 4, Part 1 asserts that, if the value function is jointly monotone, then the optimal decisions are also monotone in the system's virtual age for each belief vector $\boldsymbol{b}$. Theorem 4, Part 2 establishes that, if it is optimal to replace in a particular state, then it is optimal to replace for all larger states (the belief vector being ordered in the likelihood ratio sense). That is, within the region for which replacement is the optimal maintenance action, the preventive age replacement thresholds are monotone nondecreasing in the belief vector.

4 Numerical Examples

In this section, we illustrate our maintenance optimization framework on synthetic problem instances. We consider problems that are specifically tailored to illustrate particular properties, in addition to a large bed of randomly-parameterized problem instances. Examined are the qualitative properties of the optimal value function and resulting optimal policy. Additionally, we compare the cost of following the optimal *MOMDP* policy to several other policies.

Throughout all of our numerical examples, it is assumed that the time-to-failure, $(T|Q = y)$, follows a Weibull distribution with common shape parameter $k > 1$ and scale parameter λ_y. Additionally, it is assumed that $\lambda_1 > \lambda_2 > \cdots > \lambda_Y$. These distributional assumptions are made due to the prevalence of the Weibull distribution in modeling the time-to-failure, particularly within the context of maintenance optimization models. Additionally, under these assumptions, it is straightforward to verify that Conditions 1-3 are met.

All two-quality problem instances are coded within the MATLAB R2016a computing environment, and the three-quality problem instance is coded within the Java SE Runtime Environment 8. All codes are executed on a personal computer with a 3.50 GHz processor and 8GB of RAM.

4.1 Randomly-Generated Problem Instances

Here, 200 two-quality problems ($Y = 2$) are randomly generated with the aim of varying the problem parameters over a wide range values to assess the robustness of our *MOMDP* policy. In what follows, $U(a,b)$ denotes a continuous uniform random variable on (a,b). Fixing the number of system qualities at $Y = 2$, we randomly generate $M = 200$ problem instances. For problem $m \in \{1,\ldots,M\}$: the discount factor is denoted $\alpha^{(m)}$, where $\alpha \sim U(0.8, 0.9999)$; the cost vector is denoted $\boldsymbol{c}^{(m)} = (c_I^{(m)}, c_F^{(m)}, c_R^{(m)}, c_P^{(m)})$, where $c_I^{(m)} = 1$, $c_F^{(m)} \sim U(4,8)$, $c_R^{(m)} \sim U(1,2)$, and $c_P^{(m)} \sim c_R^{(m)} + U(1,4)$; the time-to-failure distribution shape parameter is denoted $k^{(m)}$ and scale parameter vector is denoted $\boldsymbol{\lambda}^{(m)} = (\lambda_1^{(m)}, \lambda_2^{(m)})$, where $k^{(m)} \sim U(1.1,3)$, $\lambda_2^{(m)} \sim U(1,10)$, and $\lambda_1^{(m)} \sim \lambda_2^{(m)} + U(1,10)$; the initial quality distribution is denoted $\boldsymbol{\rho}^{(m)} = (\rho_1^{(m)}, \rho_2^{(m)})$, where $\rho_1^{(m)} \sim U(0.1, 0.9)$ and $\rho_2^{(m)} = 1 - \rho_1^{(m)}$; and the inter-inspection period is denoted $\tau^{(m)}$, where $\tau^{(m)} \sim U(0.2, 1.5)$.

Because there are only two qualities, the belief state can be written as $\boldsymbol{b} = (b, 1-b)$; hence, the belief space can be simplified to the interval $[0,1]$. In order to compute the *MOMDP* policy, we discretize the interval $[0,1]$ into 1,000 states and truncate $\mathcal{X}$ to be large enough to have negligible impact on the optimal value function. When a value iteration step required the value function iterate be evaluated at a non-grid point, it is approximated using simple linear interpolation between the two nearest points. The optimal value function and policy are then obtained numerically using the value iteration algorithm.

In addition to our *MOMDP* policy, we consider three other policies: *Oracle*, *Heuristic*, and *Naive*. The *Oracle* policy is endowed with additional information in that it is given perfect information about the system quality. It then takes actions prescribed by the *MOMDP* policy but with the belief state fixed to the appropriate extreme point in Δ^Y. The *Oracle* policy provides a performance bound on the total expected discounted maintenance costs, as the additional information guarantees that, in expectation, it will outperform the *MOMDP* model. The *Heuristic* policy is determined by decoupling the problem across the belief states. Specifically, if the belief state is fixed, the problem of determining the optimal policy simplifies to solving a one-dimensional set of Bellman equations in which the time-to-failure distribution is determined by the current mixture distribution. As in the *MOMDP* policy, we begin by discretizing the interval and then solving a one-dimensional problem for each of the 1,000 belief states. This approach reduces the computational burden and storage requirements, and because each of the MDP models is monotone, we can utilize a monotone value iteration algorithm to further enhance the computational savings. The *Heuristic* policy is implemented by updating the belief about the quality of the system in the same manner as the *MOMDP* policy, but at each inspection epoch, it utilizes the one-dimensional policy corresponding to the current belief state to determine which action is taken. Finally, the *Naive* policy fixes the belief state at the initial distribution and, similar to the *Heuristic* policy, solves a one-dimensional MDP to determine whether or not to preventively maintain the system.

In order to compare the costs of the four policies, we use a simulation model. For a given simulation run, we simulate the system's time-to-failure each time it enters service, and when a system is replaced, we randomly draw a new system using the initial distribution $\boldsymbol{\rho}^{(m)}$. For each m, the simulation run length is given by the number of decision epochs $N^{(m)}$. Along each sample path, the total discounted cost is computed for each policy of interest, and these values are compared. It should be noted that, because the expected one-step costs are bounded, and the cost function is discounted, we can determine *a priori* the simulation run length needed to ensure that the total discounted cost is accurate to a fixed constant. More precisely, to guarantee that the finite approximation is within ϵ ($\epsilon > 0$) of the true total discounted cost, the number of decision epochs $N^{(m)}$ must satisfy

$$N^{(m)} \geq \frac{\ln\left[(1-\alpha^{(m)})\epsilon/C^{(m)}\right]}{\ln(\alpha^{(m)})} - 1, \quad m = 1, \dots, M,$$

where $C^{(m)}$ is any valid upper bound on the expected one-step costs. For all numerical examples, $N^{(m)}$ is chosen to correspond to $\epsilon = 0.01$ and $C^{(m)}$ is taken to be $c_I^{(m)} + c_F^{(m)} + c_P^{(m)}$.

For each problem instance $m \in \{1, \dots, M\}$, 500 sample paths are simulated and the cost of following each policy is computed. Under a particular policy and problem instance m, we denote the average total discounted maintenance cost (averaged over the 500 sample paths) by $\bar{v}_{policy}^{(m)}$, e.g., $\bar{v}_{Naive}^{(m)}$. In problem instance $m = 12$, the parameter values are as follows:

$$\begin{aligned}
\alpha^{(12)} &= 0.9904 \\
\boldsymbol{c}^{(12)} &= (1, 4.2283, 1.9530, 3.9396) \\
k^{(12)} &= 1.9516 \\
\boldsymbol{\lambda}^{(12)} &= (14.3213, 8.1914) \\
\boldsymbol{\rho}^{(12)} &= (0.4907, 0.5093) \\
\tau^{(12)} &= 0.4942
\end{aligned}$$

Problem instance $m = 12$ is noteworthy in that it exhibits the greatest discrepancy between the *MOMDP* and *Naive* policies; the *MOMDP* policy achieved a 39.47% average cost savings, i.e.,

$$\frac{\bar{v}_{Naive}^{(m)} - \bar{v}_{MOMDP}^{(m)}}{\bar{v}_{Naive}^{(m)}} \times 100\% = 39.47\%.$$

For each problem instance, we establish baseline performance by comparing each policy to the *Oracle* policy. We assess this difference by comparing the average increase in cost realized by using each policy. This increase, for a particular policy in problem instance m, is denoted by $\hat{v}_{policy}^{(m)}$. For example, when comparing the *MOMDP* policy against the *Oracle* policy in problem m, we compute

$$\hat{v}_{MOMDP}^{(m)} = \frac{\bar{v}_{MOMDP}^{(m)} - \bar{v}_{Oracle}^{(m)}}{\bar{v}_{Oracle}^{(m)}} \times 100\%.$$

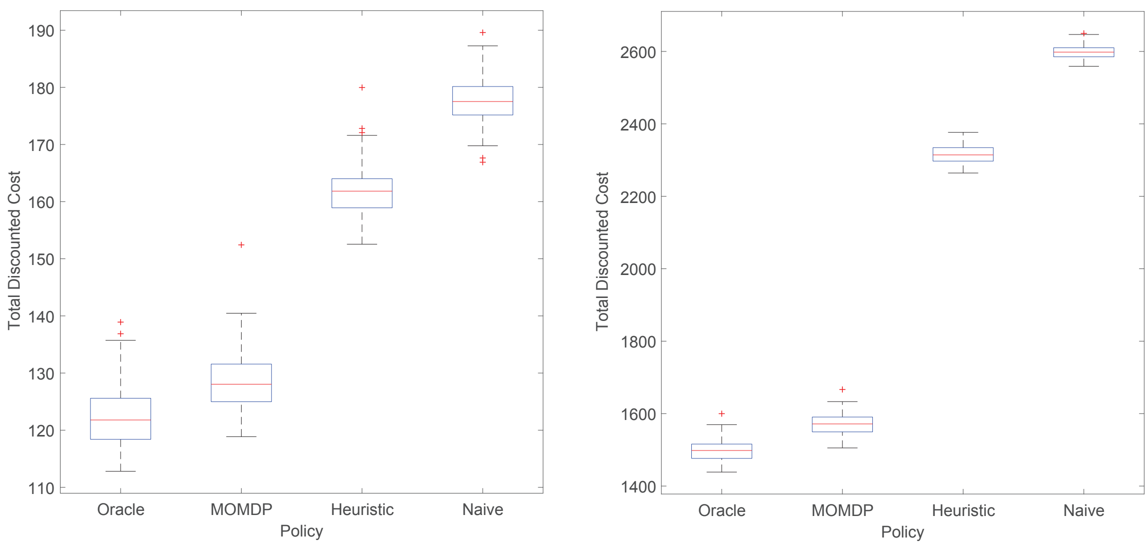
These percentage increases are then averaged over all 500 problem instances to obtain the average cost increase for a given policy, denoted by $\hat{v}_{policy}$. Table 1 summarizes the average cost increase for each policy, demonstrating significant savings achieved by utilizing the *MOMDP* policy. It is noteworthy that our model yields a nearly 20% improvement over the *Naive* policy, on average.

Table 1: Summary of policy comparison results.

$\hat{v}_{Oracle}$	$\hat{v}_{MOMDP}$	$\hat{v}_{Heuristic}$	$\hat{v}_{Naive}$
–	7.70%	20.90%	26.39%

Figure 1 depicts the cost comparison between two particular problem instances, $m = 10$ and $m = 12$. In problem instance $m = 10$, the parameter values are as follows: $\alpha^{(10)} = 0.9989$, $\boldsymbol{c}^{(10)} = (1, 7.5643, 1.7253, 4.3246)$, $k^{(10)} = 1.4123$, $\boldsymbol{\lambda}^{(10)} = (10.5496, 5.8213)$, $\boldsymbol{\rho}^{(10)} = (0.3754, 0.6246)$, and

$\tau^{(10)} = 0.9942$. In these problem instances, the difference between the performance of the *MOMDP* policy and *Oracle* policy are negligible, but there is a large performance gap between the *MOMDP* policy and the *Naive* and *Heuristic* policies. The most striking commonality between these two problem instances is that the discount factor $\alpha^{(m)}$ is large in both cases ($\alpha^{(10)}$, $\alpha^{(12)} > 0.99$). Under appropriate regularity conditions, for all discount factors sufficiently close to unity, there exists a common optimal stationary, deterministic policy that is also optimal under an average cost criterion (cf. [3, 7]). Therefore, we can expect our framework to also outperform these policies under an average cost criterion.



(a) Boxplots of total discounted cost for $m = 10$.

(b) Boxplots of total discounted cost for $m = 12$.

Figure 1: Boxplots comparing policy costs for problem instances $m = 10$ and $m = 12$.

4.2 A Specific Two-quality Problem

For the example considered in this section, the number of system qualities is again $Y = 2$. The discount rate is $\alpha = 0.99$ and the cost vector is $\boldsymbol{c} = (c_I, c_F, c_R, c_P) = (1, 2, 3, 4)$, the time-to-failure shape parameter is $k = 2$, the scale parameter vector is $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) = (12, 6)$, the initial distribution is $\boldsymbol{\rho} = (\rho_1, \rho_2) = (0.7, 0.3)$, and the inter-inspection period is $\tau = 0.2$.

To compute the *MOMDP* policy, the belief space, $[0, 1]$, is uniformly discretized into 1,000 states and $\mathcal{X}$ is truncated to be $\{0, 1, \ldots, 200, \infty\}$. The optimal value function and policy are then obtained numerically using the value iteration algorithm. When a step in the value iteration algorithm requires a value function iterate whose belief state is outside this discretization, it is approximated by linear interpolation. For each $(x, \boldsymbol{b})$ in the discretized and truncated set of states, let $v_k(x, \boldsymbol{b})$ denote the kth iterate of the value iteration algorithm. The algorithm terminates when the maximum norm of the difference between subsequent value function iterates is below 10^{-6}, that is,

$$||v_{k+1} - v_k||_\infty = \max_{x, \boldsymbol{b}} \{|v_{k+1}(x, \boldsymbol{b}) - v_k(x, \boldsymbol{b})|\} \le 10^{-6}.$$

In the case of only two qualities, the belief space is completely ordered; consequently, as seen in Figure 2, the value function exhibits monotonicity across the entire state space. In Figure 3, the *MOMDP* and *Heuristic* policies are depicted. For each fixed belief, b_1, both policies are of threshold type in age. We note that for the *MOMDP* policy, it is guaranteed to be of threshold-type by Theorem 4, Part 1. Interestingly, in the *MOMDP* policy, we see that near the interface where the repair and replacement regions meet, the age threshold is increasing in both regions. This behavior is somewhat counter intuitive as the time-to-failure is stochastically smaller (in the hazard rate sense) near this interface than it is when b_1 is nearer to 1. This behavior can be understood as a natural exploration that occurs in the *MOMDP* policy. By allowing the system to function longer near this interface, the decision maker obtains failure data that is less likely to be right-censored. This additional data can be used to increase the likelihood that a high quality system is repaired and a low quality system is replaced. Moving away from this interface (by either increasing or decreasing b_1), we see that actions become more exploitative, i.e., quickly replace systems that are likely to be low quality and allow systems that are likely to be high quality to function for longer before preventively repairing. Additionally, we see that near where $b_1 = 1$, the *MOMDP* policy and the *Heuristic* policy are nearly identical. This observation is not surprising, as Proposition 7 guarantees that they should be exactly the same when $b_1 = 1$.

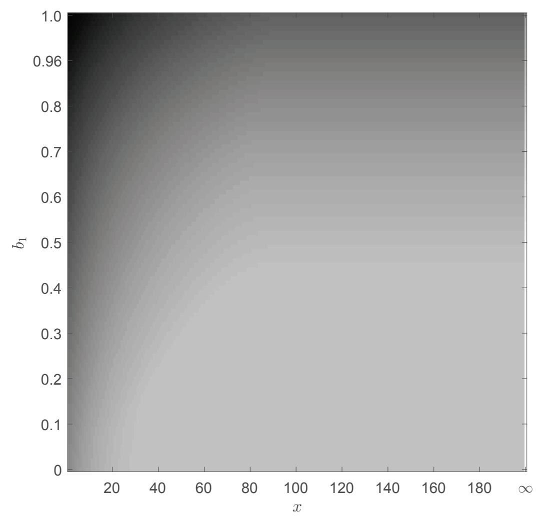
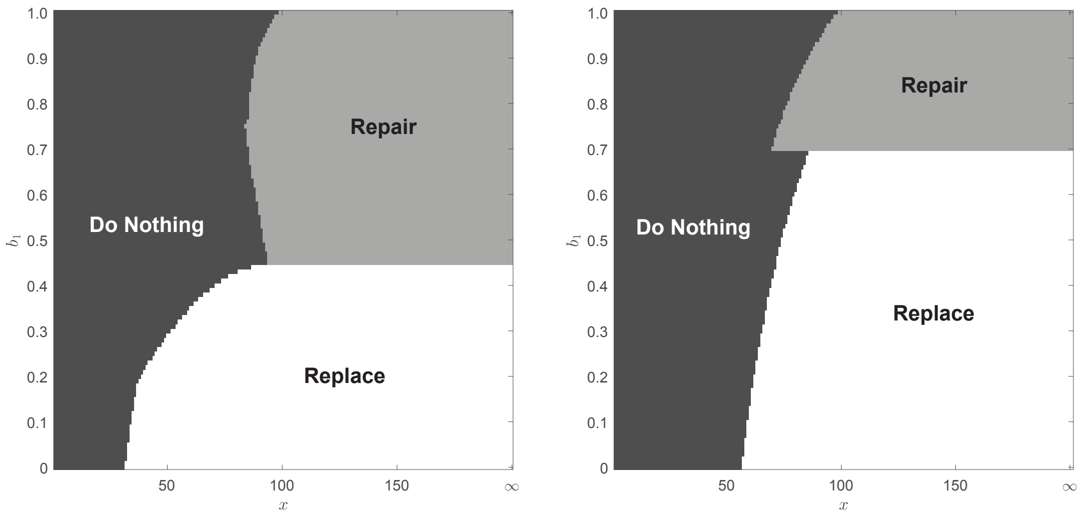


Figure 2: Depiction of the optimal value function (dark colors indicate lower costs).



(a) Depiction of the *MOMDP* policy. (b) Depiction of the *Heuristic* policy.

Figure 3: Comparison between the *MOMDP* and *Heuristic* policies.

4.3 A Specific Three-quality Problem

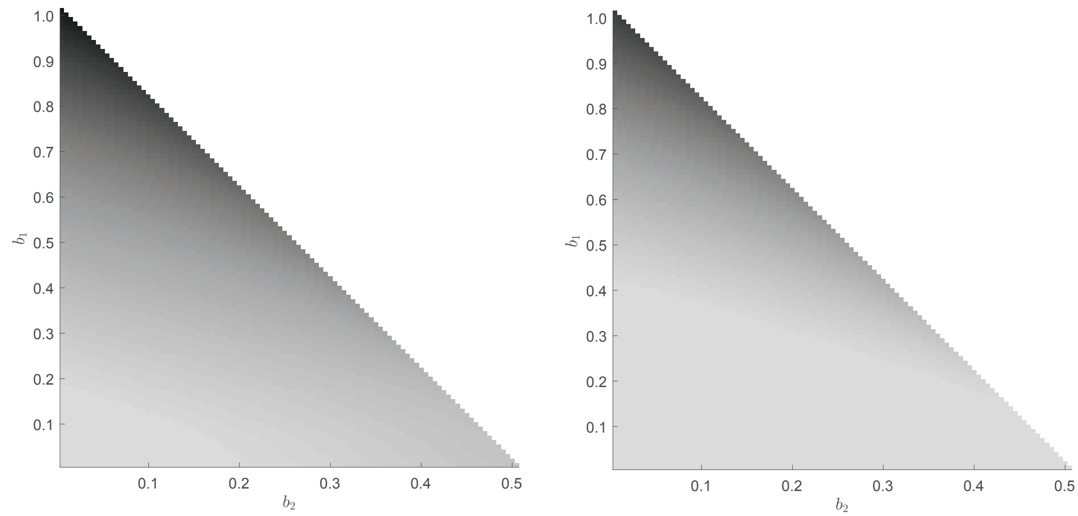
For the example considered in this section, the number of system qualities is $Y = 3$. The discount rate is $\alpha = 0.99$ and the cost vector is $\boldsymbol{c} = (c_I, c_F, c_R, c_P) = (1, 2, 3, 4)$, the time-to-failure shape parameter is $k = 2$, the scale parameter vector is $\boldsymbol{\lambda} = (\lambda_1, \lambda_2) = (12, 10, 6)$, the initial distribution is $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3) = (0.5, 0.2, 0.3)$, and the inter-inspection period is $\tau = 1$.

To compute the *MOMDP* policy, each dimension of the belief space is uniformly discretized into 500 states and $\mathcal{X}$ is truncated to be $\{0, 1, \ldots, 50, \infty\}$. The optimal value function and policy are then obtained numerically using the value iteration algorithm. When a step in the value iteration algorithm requires a value function evaluation at a belief state outside this discretization, it is approximated by bilinear interpolation (with edge cases approximated by linear or barycentric interpolation). For each $(x, \boldsymbol{b})$ in the discretized and truncated set of states, let $v_k(x, \boldsymbol{b})$ denote the kth iterate of the value iteration algorithm. The algorithm terminates when the maximum norm of the difference between subsequent value function iterates is below 10^{-6}, that is,

$$||v_{k+1} - v_k||_\infty = \max_{x,\boldsymbol{b}}\{|v_{k+1}(x, \boldsymbol{b}) - v_k(x, \boldsymbol{b})|\} \leq 10^{-6}.$$

Figure 4 depicts a portion of the optimal value function evaluated at virtual age $x = 5$ and $x = 25$. It should be noted that in each image, the origin represents the belief state $\boldsymbol{b} = \boldsymbol{e}_3 = (0, 0, 1)$, and can, therefore, be thought of as the worst belief. For this reason, we see that at $x = 5$ and $x = 25$ starting from this belief state has the highest total expected discounted cost. Similarly, starting from belief state $(1, 0, 0)$ has the lowest cost. It can also be observed that the value function exhibits monotonicity for each fixed x (as guaranteed by Theorem 1), but also across the x's, i.e., $V(5, \boldsymbol{b}) < V(25, \boldsymbol{b})$ for each $\boldsymbol{b}$.

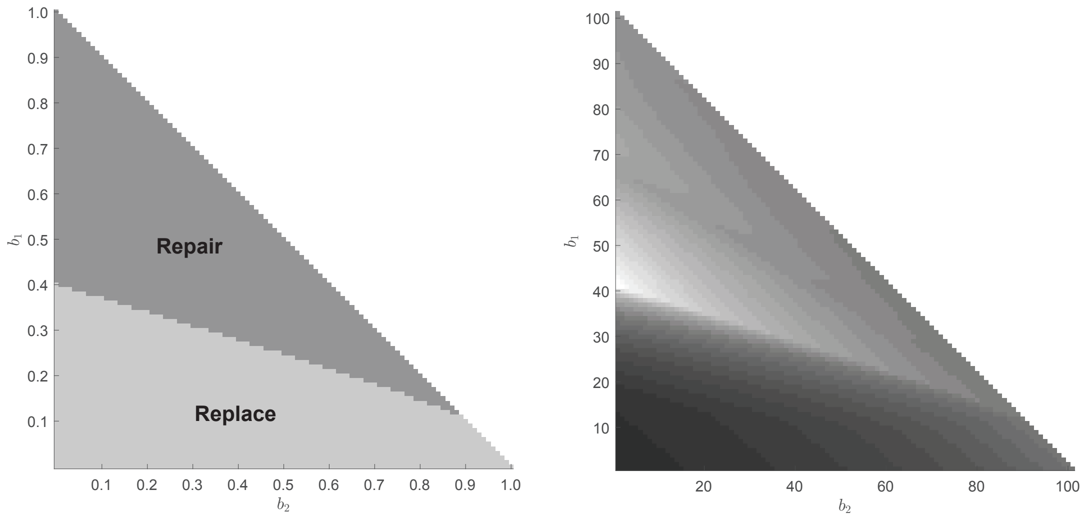
Figure 5 provides a graphical depiction of the *MOMDP* policy. By Theorem 2, the belief state space can be partitioned into two regions: one in which doing nothing or repair is optimal, and another in which doing nothing or replacement is optimal (see Figure 5(a)). It is not coincidental



(a) The optimal value function at $x = 5$. (b) The optimal value function at $x = 25$.

Figure 4: Depiction of the optimal value function (dark colors indicate lower costs).

that these regions are separated in the belief state space by a straight line; rather, it is a further consequence of Theorem 2 resulting from the partitioning of Δ^3 being related to the likelihood ratio ordering. In particular, these regions are divided in such a way that if $\boldsymbol{b} \leq_{lr} \boldsymbol{b}'$ and $\boldsymbol{b}'$ is in the repair region, then $\boldsymbol{b}$ is also in the repair region. Similarly, if $\boldsymbol{b} \leq_{lr} \boldsymbol{b}'$ and $\boldsymbol{b}$ is in the replacement region, then $\boldsymbol{b}'$ is also in the replacement region.



(a) The repair and replacement regions. (b) The preventive maintenance thresholds.

Figure 5: Depiction of the *MOMDP* policy (dark colors indicate lower thresholds).

By the joint monotonicity of the value function, and Theorem 4, Part 1, the optimal policy is guaranteed to be a threshold-type policy, for each fixed belief state. Figure 5(b) shows the age threshold for each belief state above which it is optimal to perform preventive maintenance. Unsurprisingly, the thresholds are ordered at the corner points of the plot, i.e., the threshold at $\boldsymbol{e}_3$ is the smallest and at $\boldsymbol{e}_1$ is the largest. However, in contrast to the two-quality case, the largest threshold is not when the belief state is $\boldsymbol{e}_1$, but rather at $\boldsymbol{b} = (0.4, 0, 0.6)$. Again, a ridge is formed along the interface between the repair and replacement regions where exploration is encouraged in the form of large thresholds. Additionally, we see that the thresholds are monotone in the belief state within the replacement region, but not within the repair region. By the joint monotonicity of the value function, the monotone age thresholds in the replacement region are guaranteed by Theorem 4, Part 2.

5 Conclusions and Future Work

In this work, we have considered the problem of optimally maintaining a stochastically degrading, single-unit system with heterogeneous spare parts of varying quality. To address this problem, we presented an MOMDP model and investigated its properties. Under intuitive conditions on the time-to-failure distributions, we have established monotonicity properties of the optimal value function and presented a comprehensive characterization of the optimal policy. Additionally, by way of a detailed computational study, we highlighted the cost savings that can be achieved by properly accounting for spare part heterogeneity. These numerical illustrations also revealed that the optimal policy implicitly accounts for the tradeoff between receiving high-quality, uncensored data (which improves long-term decision making) and reducing short-term maintenance costs.

The model we presented herein can be improved in a few important ways. First, our model assumes that the proportion of parts of each quality is fixed and known, and that the number of qualities and their respective time-to-failure distributions are known. Relaxing these assumptions would allow for additional model flexibility and the investigation of tradeoffs between parameter learning and maintenance decisions that exploit the current belief about the parameters. Another promising direction for future research is to relax the assumption that the inter-inspection period τ is a predetermined model parameter. Two problems related to this relaxation are worthy of further consideration: (i) determining the optimal fixed value of τ; and (ii) allowing the subsequent inter-inspection length to be set at each inspection epoch. In the latter problem, shorter inter-inspection intervals would provide higher-quality information but with an increase in cost. Due to this tradeoff, we suspect the optimal policy may be difficult to fully characterize.

Acknowledgements. We are grateful to two anonymous referees and the Associate Editor for their helpful comments and suggestions.

References

[1] R. E. Barlow and F. Proschan. *Mathematical Theory of Reliability.* John Wiley & Sons, Inc., New York, NY, 1965.

[2] L. Bian and N. Gebraeel. Computing and updating the first-passage time distribution for randomly evolving degradation signals. *IIE Transactions*, 44(11):974–987, 2012.

[3] D. Blackwell. Discrete dynamic programming. *Annals of Mathematical Statistics*, 33(2):719–726, 1962.

[4] J. H. Cha and M. Finkelstein. Stochastic analysis of preventive maintenance in heterogeneous populations. *Operations Research Letters*, 40(5):416–421, 2012.

[5] N. Chen, Z. Ye, Y. Xiang, and L. Zhang. Condition-based maintenance using the inverse Gaussian degradation model. *European Journal of Operational Research*, 243(1):190–199, 2015.

[6] R. Dekker. Applications of maintenance optimization models: A review and analysis. *Reliability Engineering & System Safety*, 51(3):229–240, 1996.

[7] C. Derman. On sequential decisions and Markov chains. *Management Science*, 9(1):16–24, 1962.

[8] A. H. Elwany, N. Z. Gebraeel, and L. M. Maillart. Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research*, 59(3):684–695, 2011.

[9] A. L. Hartzell, M. G. da Silva, and H. R. Shea. *MEMS Reliability.* MEMS Reference Shelf. Springer, Boston, MA, 2011.

[10] F. Jensen and N. E. Petersen. *Burn-In: An Engineering Approach to the Design and Analysis of Burn-in Procedures.* John Wiley & Sons Ltd., New York, NY, 1982.

[11] W. Kuo and Y. Kuo. Facing the headaches of early failures: A state-of-the-art review of burn-in decisions. *Proceedings of the IEEE*, 71(11):1257–1266, 1983.

[12] J. Lawless and M. Crowder. Covariates and random effects in a Gamma process model with applications to degradation and failure. *Lifetime Data Analysis*, 10(3):213–227, 2004.

[13] J. J. McCall. Maintenance policies for stochastically failing equipment: A survey. *Management Science*, 11(5):493–524, 1965.

[14] J. Mi. Burn-in and maintenance policies. *Advances in Applied Probability*, 26(1):207–221, 1994.

[15] J. Mi. Minimizing some cost functions related to both burn-in and field use. *Operations Research*, 44(3):497–500, 1996.

[16] S. C. W. Ong, S. W. Png, D. Hsu, and W. S. Lee. Planning under uncertainty for robotic tasks with mixed observability. *The International Journal of Robotics Research*, 29(8):1053–1068, 2010.

[17] C.-Y. Peng and S.-T. Tseng. Mis-specification analysis of linear degradation models. *IEEE Transactions on Reliability*, 58(3):444–455, 2009.

[18] H. Pham and H. Wang. Imperfect maintenance. *European Journal of Operational Research*, 94(3):425–438, 1996.

[19] W. P. Pierskalla and J. A. Voelker. A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, 23(3):353–388, 1976.

[20] M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer, New York, NY, 2007.

[21] A. Sharma, G. Yadava, and S. Deshmukh. A literature review and future perspectives on maintenance optimization. *Journal of Quality in Maintenance Engineering*, 17(1):5–25, 2011.

[22] Y. Sherif and M. Smith. Optimal maintenance models for systems subject to failure – A review. *Naval Research Logistics Quarterly*, 28(1):47–74, 1981.

[23] N. D. Singpurwalla. Survival in dynamic environments. *Statistical Science*, 10(1):86–103, 1995.

[24] C.-C. Tsai, S.-T. Tseng, and N. Balakrishnan. Optimal design for degradation tests based on Gamma processes with random effects. *IEEE Transactions on Reliability*, 61(2):604–613, 2012.

[25] C. Valdez-Flores and R. M. Feldman. A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics*, 36(4):419–446, 1989.

[26] C. van Oosterom, H. Peng, and G. van Houtum. Maintenance optimization for a Markovian deteriorating system with population heterogeneity. *IISE Transactions*, 49(1):96–109, 2017.

[27] H. Wang. A survey of maintenance policies of deteriorating systems. *European Journal of Operational Research*, 139(3):469–489, 2002.

[28] W. Wang. A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance. *International Journal of Production Research*, 38(6):1425–1436, 2000.

[29] X. Wang. Wiener processes with random effects for degradation data. *Journal of Multivariate Analysis*, 101(2):340–351, 2010.

[30] Y. Xiang, D. W. Coit, and Q. Feng. n Subpopulations experiencing stochastic degradation: reliability modeling, burn-in, and preventive replacement optimization. *IIE Transactions*, 45(4):391–408, 2013.

[31] Y. Xiang, D. W. Coit, and Q. Feng. Accelerated burn-in and condition-based maintenance for n-subpopulations subject to stochastic degradation. *IIE Transactions*, 46(10):1093–1106, 2014.

[32] Z. Ye, Y. Shen, and M. Xie. Degradation-based burn-in with preventive maintenance. *European Journal of Operational Research*, 221(2):360–367, 2012.

[33] M. Zhang, Z. Ye, and M. Xie. A condition-based maintenance strategy for heterogeneous populations. *Computers & Industrial Engineering*, 77:103–114, 2014.

Appendix: Proofs of Lemmas

Proof of Lemma 1

Proof. Define $\lim_{x\to\infty} B_y(x, \boldsymbol{b}) = \ell_y$, then if ℓ_y exists for ever y, $B(x, \boldsymbol{b}) \to (\ell_1, \ldots, \ell_Y)$. Now, by definition,

$$B_y(x, \boldsymbol{b}) = \frac{g(x, y)}{G(x, \boldsymbol{b})} b_y,$$

where clearly, by Condition 2, we have that $g(x, y) \to 1$ and $G(x, \boldsymbol{b}) = \sum_m g(x, m) b_m \to \sum_m b_m = 1$. Therefore, $B_y(x, \boldsymbol{b}) \to b_y$ and $B(x, \boldsymbol{b}) \to \boldsymbol{b}$. Similarly, we consider $\bar{B}_y(x, \boldsymbol{b})$, where, after some algebraic manipulation,

$$\begin{aligned}
\bar{B}_y(x, \boldsymbol{b}) &= \frac{\bar{g}(x, y) b_y}{\sum_m \bar{g}(x, m) b_m} \\
&= \frac{b_y}{\sum_m \frac{\bar{g}(x,m)}{\bar{g}(x,y)} b_m} \\
&= \frac{b_y}{b_y + \sum_{m \neq y} \frac{\bar{g}(x,m)}{\bar{g}(x,y)} b_m}.
\end{aligned} \tag{26}$$

For $y = 1$, by Condition 3, the expression in (26) converges to 1. ∎

Proof of Lemma 2

Proof. In what follows, the norm $||\cdot||$ will denote the Euclidean norm. For $\epsilon > 0$, consider $\delta(\epsilon) = \epsilon/Y$. Then, for any $\boldsymbol{b}, \boldsymbol{b}' \in \Delta^Y$ such that $||\boldsymbol{b} - \boldsymbol{b}'|| < \delta(\epsilon)$ we seek to show that $\left|\bar{G}(x, \boldsymbol{b}) - \bar{G}(x, \boldsymbol{b}')\right| < \epsilon$. First, we note that $||\boldsymbol{b} - \boldsymbol{b}'|| < \delta(\epsilon)$ implies $\left|b_y - b'_y\right| < \delta(\epsilon)$ for all $y \in \mathcal{Y}$. Now, by the non-negativity of $\bar{g}(x, y)$ and the triangle inequality, we know that

$$\left|\bar{G}(x, \boldsymbol{b}) - \bar{G}(x, \boldsymbol{b}')\right| = \left|\sum_y \bar{g}(x, y)(b_y - b'_y)\right| \leq \sum_y \bar{g}(x, y) \left|b_y - b'_y\right|,$$

but by the bound on $\left|b_y - b'_y\right|$, we have that

$$\left|\bar{G}(x, \boldsymbol{b}) - \bar{G}(x, \boldsymbol{b}')\right| < \frac{\epsilon}{Y} \sum_y \bar{g}(x, y) \leq \frac{\epsilon}{Y} \sum_y 1 = \epsilon.$$

Therefore, $\bar{G}(x, \boldsymbol{b})$ is continuous in $\boldsymbol{b}$, and because $G(x, \boldsymbol{b}) = 1 - \bar{G}(x, \boldsymbol{b})$, we conclude that $G(x, \boldsymbol{b})$ is likewise continuous in $\boldsymbol{b}$.

Next, note that the function $\bar{B}(x, \cdot)$ is continuous if, and only if, $\bar{B}_y(x, \cdot) : \mathbb{R}^Y \to \mathbb{R}$ is continuous for all $y \in \mathcal{Y}$. By the continuity of $\bar{G}$, we note that $\bar{B}_y(x, \boldsymbol{b}) = \left(\bar{G}(x, \boldsymbol{b})\right)^{-1} \bar{g}(x, y) b_y$ is a product of continuous functions and is, therefore, continuous. Hence, by the continuity of its components, $\bar{B}(x, \cdot)$ is continuous. The proof that B is continuous is similar. ∎

Proof of Lemma 3

Proof. By induction, if $v^0(x, \boldsymbol{b}) = 0$ for all $(x, \boldsymbol{b}) \in \mathcal{X} \times \Delta^Y$, then

$$v^1(x, \boldsymbol{b}) = \begin{cases} c_I + c_R + c_F, & x = \infty, \\ c_I, & x < \infty, \end{cases}$$

so the base case holds. Now, assume $v^k(x, \boldsymbol{b})$ is continuous in $\boldsymbol{b}$ and note that

$$v^{k+1}(\infty, \boldsymbol{b}) = c_I + c_F + \min\{c_P + \alpha v^k(0, \boldsymbol{\rho}),\, c_R + \alpha v^k(0, \boldsymbol{b})\}.$$

Because the minimum of continuous functions is again continuous, we conclude that $v^{k+1}(\infty, \boldsymbol{b})$ is continuous in $\boldsymbol{b}$. Now, for finite $x \in \mathcal{X}$,

$$v^{k+1} = \min\{v^{k+1}(\infty, \boldsymbol{b}) - c_F,\, C^k_{DN}(x, \boldsymbol{b})\},$$

where $C^k_{DN}(x, \boldsymbol{b}) = c_I + \alpha(\bar{G}(x, \boldsymbol{b}) v^k(x+1, \bar{B}(x+1, \boldsymbol{b})) + G(x, \boldsymbol{b}) v^k(\infty, B(x+1, \boldsymbol{b})))$. Therefore, we proceed to show that $C^k_{DN}(x, \boldsymbol{b})$ is continuous to complete the proof. By Lemma 2, it is known that $\bar{B}$ is continuous; hence,

$$\lim_{\boldsymbol{b}_n \to \boldsymbol{b}} \bar{B}(x+1, \boldsymbol{b}_n) = \bar{B}(x+1, \boldsymbol{b}).$$

Then, by the induction hypothesis, we see that

$$\lim_{\boldsymbol{b}_n \to \boldsymbol{b}} v^k(x+1, \bar{B}(x+1, \boldsymbol{b}_n)) = v^k(x+1, \lim_{\boldsymbol{b}_n \to \boldsymbol{b}} \bar{B}(x+1, \boldsymbol{b}_n)) = v^k(x+1, \bar{B}(x+1, \boldsymbol{b}));$$

thus, $v^k(x+1, \bar{B}(x+1, \boldsymbol{b}_n))$ is continuous. Similarly, $v^k(\infty, B(x+1, \boldsymbol{b}_n))$ is also continuous. Lastly, we note that G and $\bar{G}$ are continuous by Lemma 2. Because $C^k_{DN}(x, \boldsymbol{b})$ is the composition of continuous functions, it is also continuous. ■

Proof of Lemma 4

Proof. We first show that, for all $x < \infty$, $V(x, \boldsymbol{b})$ is finitely bounded. By (2), we note that $V(0, \boldsymbol{b}) \leq c_I + c_R + \alpha V(0, \boldsymbol{b})$. Rearranging terms, observe that $V(0, \boldsymbol{b}) \leq (c_I + c_R)/(1 - \alpha)$. For any $x < \infty$, $V(x, \boldsymbol{b}) \leq c_I + c_R + \alpha V(0, \boldsymbol{b}) \leq c_I + c_R + \alpha(c_I + c_R)/(1 - \alpha)$. By Condition 2 and this finiteness, we have that

$$\lim_{x \to \infty} \bar{G}(x, \boldsymbol{b}) V(x+1, \bar{B}(x+1, \boldsymbol{b})) = 0.$$

By Condition 2, Lemma 1, and Lemma 3,

$$\lim_{x \to \infty} G(x, \boldsymbol{b}) V(\infty, B(x+1, \boldsymbol{b})) = V(\infty, \lim_{x \to \infty} B(x+1, \boldsymbol{b})) = V(\infty, \boldsymbol{b}).$$

Therefore, $V_{DN}(x, \boldsymbol{b}) \to c_I + \alpha V(\infty, \boldsymbol{b})$. ■