Temporal and Spatial Considerations in Maintenance Planning

by

Shadi Sanoubar

B.S., Sharif University of Technology, 2016

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Shadi Sanoubar

It was defended on

June 9, 2022

and approved by

Lisa M. Maillart, Ph.D., Professor, Department of Industrial Engineering

Oleg A. Prokopyev, Ph.D., Professor, Department of Industrial Engineering

Jayant Rajgopal, Ph.D., Professor, Department of Industrial Engineering

Jeffrey P. Kharoufeh, Ph.D., Professor, Department of Industrial Engineering, Clemson University

Hyo Kyung Lee, Ph.D., Assistant Professor, School of Industrial Management Engineering, Korea University

Dissertation Directors:

Lisa M. Maillart, Ph.D., Professor, Department of Industrial Engineering

Oleg A. Prokopyev, Ph.D., Professor, Department of Industrial Engineering

Temporal and Spatial Considerations in Maintenance Planning

Shadi Sanoubar, PhD

University of Pittsburgh, 2022

Maintenance spending is well-known to constitute a substantial part of total production and service costs. We focus on optimal planning of maintenance activities in several novel settings. In each setting, we formulate a mathematical optimization model using stochastic modeling techniques and establish the structural properties of the optimal policy through theoretical derivations. We provide additional policy insights using numerical observations and develop easy-to-implement and high-performing heuristic policies.

Specifically, we first study an age-replacement setting (with minimal repair) in which the maintenance worker may be unpunctual. That is, the actual preventive replacement times may deviate from the prescribed replacement times in a probabilistic manner. We formulate a long-run expected cost-rate minimization model and compare the optimal solution and its performance to those when the unpunctual behavior is assumed to be either absent or independent of the prescribed replacement time.

Next, we consider an age-replacement setting (without minimal repair) in which replacement costs are non-decreasing in system age. This assumption is motivated by factors such as decreasing salvage value or increasing costs associated with obtaining spare parts. We formulate a long-run expected cost-rate minimization model that captures this dependency and compare the optimal solution and its performance to those for the case in which replacement costs are assumed to be constant.

Finally, we consider the problem of performing condition-based maintenance on a set of geographically distributed assets via a single maintenance resource that travels between the assets' locations. We use a graph representation to model possible geographical locations of the resource, including idling and asset locations and the links between them. We formulate a Markov decision process to dynamically obtain the optimal positioning of the maintenance resource and the optimal timing of the interventions that the resource performs.

Keywords: Maintenance Optimization, Stochastic Processes, Markov Decision Process.

Table of Contents

Preface ix

1.0 Introduction 1

1.1 Time-Based Maintenance 1

1.2 Condition-Based Maintenance 4

2.0 Optimal Age-Replacement under Time-Dependent Unpunctual Policy Implementation 6

2.1 Introduction 6

2.2 Model Formulation 8

2.3 Existence of a Unique Optimal Solution 15

2.4 Analytical Results 17

2.4.1 Analysis of $\Omega(T^*)$ vs. $\hat{\Omega}(\hat{T}^*)$ vs. $\widetilde{\Omega}(\widetilde{T}^*)$ 18

2.4.2 Analysis of T^* vs. $\hat{T}^*$ vs. $\widetilde{T}^*$ 21

2.5 Heuristic Policies and Cost-rate Bounds 26

2.5.1 Heuristic Policies 26

2.5.2 Cost-rate Bounds 28

2.6 A Numerical Example for $z(\hat{T}^*) \neq 1$ & $z(T^*) \neq 1$ 35

2.7 Concluding Remarks 36

3.0 Optimal Age-Replacement under Time-Dependent Replacement Costs 39

3.1 Introduction 39

3.2 Model Formulation 40

3.3 Existence of a Unique Optimal Solution 43

3.4 Increasing vs. Constant Replacement Costs 45

3.4.1 $\widetilde{\Omega}(\widetilde{T}^*)$ vs. $\Omega(T^*)$ and $\widetilde{T}^*$ vs. T^* 46

3.4.2 Price of Ignoring Age-dependency 50

3.5 Non-Instantaneous Replacements 52

3.6 Concluding Remarks 60

4.0 Optimal Condition-Based Maintenance via a Mobile Maintenance Resource . . . 61
4.1 Introduction . . . 61
4.2 Model Formulation . . . 64
4.3 Structural Properties . . . 69
4.4 Policy Insights . . . 71
4.4.1 Maintenance Thresholds . . . 71
4.4.2 Positioning and Deterioration Conditions . . . 75
4.4.3 Idling and Graph Centrality . . . 78
4.5 Additional Performance Metrics . . . 83
4.6 Heuristic Policies . . . 88
4.6.1 Heuristic Policy Definitions . . . 88
4.6.2 Optimality Gap . . . 92
4.6.3 Large Problem Instances . . . 94
4.6.3.1 Other Approaches for Obtaining Anticipatory Idling Nodes . . . 96
4.6.3.2 Modifying the Repair Threshold . . . 97
4.7 Concluding Remarks . . . 99
5.0 Summary of Contributions and Future Work . . . 105
Appendix A. Chapter 2.0: Proofs . . . 108
Appendix B. Chapter 3.0: Proofs . . . 114
Appendix C. Chapter 4.0: Proofs . . . 119
Appendix D. Chapter 4.0: Control-Limit Rule Violation . . . 124
Appendix E. Chapter 4.0: Downtime and Graph Centrality . . . 126
Appendix F. Chapter 4.0: Deterioration Probability Matrices of Section 4.6 129
Bibliography . . . 131

List of Tables

1 Summary of age replacement models with minimal repair 14

2 Numerical example of the bounds in Propositions 6-8 for decreasing $z(T)$ 31

3 Numerical examples of the bounds in Propositions 6-8 for increasing $z(T)$ . . . 32

4 Optimal cost-rate and replacement time for $z(\hat{T}^*) \neq 1$ or $z(T^*) \neq 1$ 36

5 Percent increase in long-run cost-rate induced by suboptimal policies 52

6 Case of $\beta_p < \beta_r \ll \mu$ 56

7 Case of $\beta_p < \beta_r \leq \mu$ 58

8 Summary of findings on $\mathcal{C}(T)$ and $\mathcal{A}(T)$ 59

9 Summary of the related literature and our contributions 64

10 Performance metrics for a (2,4)-banana tree 85

11 Performance metrics for a (3,3)-grid graph 86

12 Heuristic policy performance against the optimal policy 93

13 Heuristic policy performance under $\bar{\mu}^*$ in tree graph instances 104

14 Heuristic policy performance under $\bar{\mu}^*$ in grid graph instances 104

List of Figures

1 Possible cycle dynamics under age replacement with minimal repair 11

2 Long-run cost-rate for Example 1 16

3 Long-run cost-rates for Example 2 20

4 Long-run cost-rate for Example 3 22

5 Preferred heuristic replacement age under different parameter settings 28

6 Example sequence of replacement times for a given age-replacement policy . . . 43

7 Long-run cost-rate for Example 8 45

8 Long-run cost-rate for Example 9 49

9 Long-run cost-rates for Example 9 50

10 Long-run cost-rates for Example 11 55

11 Long-run cost-rates for case of $\beta_p < \beta_r \leq \mu$ 57

12 Example of a graph with four assets 66

13 An excerpt of the optimal policy for Example 13 74

14 An excerpt of the optimal policy for Example 14 75

15 An excerpt of the optimal policy for Example 15 75

16 An excerpt of the optimal policy for unbalanced deterioration levels 76

17 Plot of value function and the corresponding optimal actions 77

18 Heat maps for Example 16 79

19 Heat maps for Example 17 82

20 Instances of graphs used in the numerical study of Section 4.6.3 95

21 Heuristic policy performance in large tree graph instances 101

22 Heuristic policy performance in large grid graph instances 102

23 Heuristic policy performance under different approaches of obtaining set A . . . 103

24 Excerpt of the optimal policies for the examples in Appendix D 125

25 Average percentage of downtime for six assets in different graph configurations . 127

26 Plot of average percentage of downtime against the closeness centrality 128

Preface

I dedicate this dissertation to my parents, Naiyer and Naser.

Thank you to my advisors, Dr. Lisa Maillart and Dr. Oleg Prokopyev, for the countless hours spent providing feedback, for guiding me through the research process, for teaching me how to organize my writings, for your advice and great ideas. You always acknowledged my accomplishments, encouraged me, and believed in me even when I did not.

Sincere appreciation to my committee members, Dr. Kharoufeh, Dr. Lee, and Dr. Rajgopal for all the guidance and support, for taking the time to write my recommendation letters, attending my INFORMS talks, and insightful conversations. Thank you to Dr. Kai He and Dr. Bram de Jonge for your collaboration and helping me prepare the second and fourth chapters of this dissertation.

My gratitude to my husband, Alireza, who encouraged me to pursue my doctoral degree in the US, and supported me in every way possible.

Thank you to Dr. Kirk Pruhs for making me and Alireza try new beers. I enjoyed watching El Classico games with you and have learned so much from our conversations.

I am grateful for the friends who have kept me energized and on track, including Moataz Abdulhafez, Moataz Elsisy, Sajad Haghanifar, Anna Svirsko, Arnab Bhattacharya, David Abdulmalak, Hanie Eskandari, Rana Shariat, as well as everyone in the INFORMS student chapter and IE Grad soccer squad. All my best to my soul sister, Jing Yang; I feel so lucky to have a loving and caring friend like you.

To my sister, Mahsa: thank you for being a ray of sunshine. Your persistence, hard work, and energy has been inspiring to me.

To Afkham, Vida, Mozhgan, Furugh, Asal, MohammadReza, Hushang, Farhang, Hamide and Mahnaz: thank you for the laughter, distraction, and love.

Acknowledgements. The content in Chapter 2 is partly reproduced from the following manuscript: S. Sanoubar, K. He, L.M. Maillart, O.A. Prokopyev, Optimal Age-Replacement in Anticipation of Time-Dependent, Unpunctual Policy Implementation, *IEEE Transactions on Reliability*, Vol. 70/3 (2021), pp. 1177-1192. This work was supported in part by the National Science Foundation under Grant CMMI-1400009 and in part by the Air Force Office of Scientific Research under Grant FA2386-12-1-3032. Moreover, the content in Chapter 3 is partly reproduced from the following manuscript: S. Sanoubar, L.M. Maillart, O.A. Prokopyev, Age Replacement Policies Under Age-Dependent Replacement Costs, *IISE Transactions*, Vol. 53/4 (2021), pp. 425-436. Lastly, the content in Chapter 4 is supported by the University of Pittsburgh Central Research Development Fund Seeding Grant.

1.0 Introduction

Maintenance spending is well-known to constitute a substantial part of the operating budgets in organizations with large investments in machinery and equipment [113]. Hence, research on maintenance optimization has gained tremendous attention since the 1960's [47]. Recent maintenance practices involve planning maintenance activities such as measurements, adjustments, and replacements *before* asset/equipment failure. Proper planning and implementation of these maintenance tasks can reduce total failure costs and equipment downtime.

Maintenance strategies are often classified into time-based and condition-based maintenance [50, 54]. In *time-based maintenance*, the decisions are determined based on the equipment time to failure distribution; whereas in *condition-based maintenance*, these decisions are based on the information collected through a condition monitoring process.

Time-based maintenance is typically easy-to-implement as only system age or service time has to be recorded. However, under this strategy, remaining life of systems may be wasted if they are still in *good* conditions when a *preventive* maintenance is performed. *Preventive* maintenance is referred to maintaining an asset before failure in a time-based maintenance setting [3].

On the other hand, condition-based maintenance can prescribe maintenance interventions *when needed*, resulting in more cost-effective strategies. That being said, condition-based maintenance should only be adopted if its benefits outweigh the costs associated with implementing this strategy. These costs include installing condition-monitoring equipment and software, and training/hiring experienced personnel. Therefore, both strategies are widely adopted across many industries [3, 46, 58, 115].

1.1 Time-Based Maintenance

Age-replacement is one of the most common time-based maintenance policies for systems with increasing failure rates [131], and continues to receive attention in the literature [11, 137, 138]. Under an age-replacement policy, a system is replaced at some planned age T (preventive maintenance) or is maintained at failure (reactive maintenance), whichever occurs first. When the system fails before age T, it is either minimally repaired (i.e., has the same failure rate as before failure after minimal repair) or replaced (i.e., the system is restored to as-good-as-new.)

Unrealistic assumptions underlying time-based maintenance strategies is cited as one of the practical issues [3]. For example, it is often assumed that maintenance actions are carried out on time, or the cost of maintenance actions are independent of equipment age or price at the time of replacement. Even though these assumptions simplify the analysis and computation of the optimal maintenance policies, they may perform poorly in practice.

In many settings, the specification of maintenance policies and their implementation are carried out by separate parties [43, 102]. Therefore, improper policy implementation, e.g., unpunctual maintenance, can occur, leading to poor system performance. For instance, when maintaining hydraulic accumulators, early or late deviations from the pre-charge maintenance interval are common and can be costly [34]. Moreover, according to a study by engineering firm SWECO, delayed maintenance on bridges and tunnels can lead to major traffic issues [101]. In fact, an important maintenance metric used by organizations is Preventive Maintenance Compliance, which is defined as the percentage of preventive maintenance activities completed within a predefined time window [41]. To mitigate the loss associated with such deviations, potential implementation errors may be anticipated by planners during policy specification.

Chapter 2 challenges the assumption of timely implementation of maintenance actions in an age-replacement framework. Specifically, it assumes that the actual time of preventive replacements that are carried out by a maintenance worker may deviate from the prescribed replacement times in a probabilistic manner, and the degree of deviation depends on the prescribed replacement time. We focus on modeling the temporal impact of the prescribed replacement policy on the worker's unpunctual behavior and analyzing its effect on the optimal policy and its performance. We generalize a previous work in which the degree of deviation is assumed to be independent of the prescribed replacement time [64].

Another limiting assumption in most studies on age-replacement policies is considering stationary maintenance costs [12, 13, 138]; in practice, however, these costs can depend on system age. There are multiple factors that contribute to age-dependent maintenance costs. Decreasing salvage value of equipment and system components is one such factor [16, 39, 116]. For example, preventive replacement of computers is a widely adopted policy for medium to large size businesses that have low tolerance for downtime [84, 108]. In these businesses, adjusting replacement ages by considering salvage values can significantly reduce maintenance expenditures. Similarly, in construction and mining companies, salvage value is one of the most important factors that impacts decisions such as repairing, replacing or disposing of equipment [52]; indeed, various studies aim to model the residual value of heavy construction equipment [40, 86, 103]. The study in [53] assumes decreasing salvage value in deriving dynamic replacement policies for an agency that maintains large fleets of vehicles and specialized equipment. The case of decreasing salvage value can also arise in applications where the equipment is refurbished for reuse, and refurbishment costs increase in equipment age. For instance, utility distribution transformers are often refurbished, and their refurbishment may involve simple adjustments, additions of extra components, or a complete rewind [14].

Moreover, in many settings (e.g., medical equipment, switchgears in power systems, computer servers and IT infrastructures), sourcing replacement parts can become more challenging as equipment ages because they are either no longer available or become more expensive [7, 74, 81, 89, 109]. Finally, sourcing trained technicians may also become more challenging as equipment ages, resulting in increasing replacement costs [109].

In view of the discussion above, Chapter 3 challenges the assumption of constant replacement costs in an age-replacement framework. In particular, we formulate a long-run expected cost-rate minimization model with instantaneous replacements that captures this dependency, and provide conditions under which there exists a unique optimal solution. We provide analytical and numerical results that compare the cost-rate minimizing optimal replacement policy, and its performance, to those for the case in which replacement costs are assumed to be constant. We also consider non-instantaneous replacements, and compare cost-rate minimizing and availability maximizing policies.

1.2 Condition-Based Maintenance

Recent advances in sensor technologies and remote monitoring tools have facilitated real-time tracking of asset health parameters that enable the implementation of condition-based maintenance (CBM) strategies. A CBM strategy prescribes maintenance activities dynamically/adaptively over time based on condition monitoring information, resulting in more cost-effective maintenance plans compared to traditional approaches. The global condition monitoring market size is projected to grow from $2.6 billion (USD) in 2021 to $3.6 billion (USD) by 2026, with North America as a key market for these technologies [88]. This growth in the condition monitoring market is also fueled by recent shifts in many industries toward automation, spurred in part by the "great resignation," which resulted in an increase of up to 37% in robotic orders of North American companies in 2021 compared to 2020 [93]. Numerous industries, especially those with capital-intensive investments, have recently adopted CBM strategies. Some examples include transportation networks, marine technologies, aerospace and defense industries, oil and gas pipelines, and IT infrastructures [92, 105, 124, 135].

Common characteristics shared by many of these industry applications include a set of geographically dispersed assets that must be maintained by a limited number of maintenance resources. In such settings, the maintenance resources may be positioned between the asset locations and travel to the assets to maintain (e.g., repair or replace) them.

The maintenance resources in these settings may be human crews, heavy equipment and materials, or, increasingly, self-propelled maintenance robots. Consider, for example, railway transportation, in which crews and equipment are positioned in anticipation of performing many types of activities to maintain its infrastructure composed of tracks, bridges, tunnels, signals, and other equipment [66]. In terms of inspection and maintenance robots, it is worth noting that their market size is projected to grow from $1.7 billion (USD) in 2020 to $3.5 billion (USD) by 2028 [44, 94]. Example applications of these robots include customer fulfillment centers, subsea installations and technologies, water pipelines, and computer server centers [4, 10, 26, 83].

In all of these applications, an effective maintenance plan requires the integrated optimization of two decision-making processes, namely, the *timing of maintenance interventions* and the *repositioning of maintenance resources* (e.g., specialized equipment, human crews, and robots) using the assets' *condition information*. This complex decision space gives rise to unique and as of yet, not well understood, trade-offs. For example, when maintaining a set of geographically dispersed condition-monitored assets, it may be optimal to maintain an asset earlier than it would be for that asset in isolation if a maintenance resource is currently "sufficiently close" to the location of the asset. Or, based on current condition information it may be optimal to *reposition* maintenance resources or have them *idle* in key locations in anticipation of asset deterioration that would prompt a future maintenance action.

The literature on condition-based maintenance mainly focuses on optimal timing of maintenance actions [37, 71, 126]. Integrating positioning decisions with maintenance timing decisions, however, has not been adequately studied in the literature. A few studies integrate spatial considerations with preventive maintenance planning for time-based maintenance [69, 85], however, they do not address condition-based maintenance.

The purpose of Chapter 4 is to integrate positioning and maintenance decisions under a condition-based maintenance framework, and study the novel trade-offs associated with such settings. In particular, we formulate a Markov decision process to dynamically obtain the optimal positioning of the maintenance resource and the optimal timing of the interventions that the resource performs. These decisions are made as a function of the conditions of the assets and the current location of the maintenance resource to minimize total expected costs, which include downtime, travel, and maintenance expenses. We develop insights on properties of the optimal policies (analytically and numerically) and how they are affected by the graph structure.

2.0 Optimal Age-Replacement under Time-Dependent Unpunctual Policy Implementation

2.1 Introduction

In this chapter, we consider a setting where a maintenance *planner* prescribes an age-replacement policy for a degrading system and a maintenance *worker* implements the policy. The maintenance worker may perform replacement earlier or later than intended and the magnitude of this potential deviation is captured through the random variable Y. Unpunctual maintenance-policy implementation is studied in [64] and [132]. The authors in [64] provide analytical results based on the assumption that this deviation Y is independent of the prescribed replacement time. In other words, no matter how near or far into the future replacement is scheduled, the unpunctual tendencies of the maintenance worker remain unchanged. The authors in [132] consider a different setting where a free repair warranty expires when either the item age or the total usage exceeds a certain limit, whichever occurs first, but like the model in [64], assume that the potential deviation from the intended policy is independent of the policy set by the maintenance planner.

In practice, however, adherence to a plan and the degree of unpunctuality can be influenced by how far into the future activities are scheduled. This phenomenon is well-documented in the appointment scheduling literature. For instance, in [133], the authors examine the relationship between the rate of missed and canceled medical appointments and the scheduling interval, i.e., the amount of time between when an appointment is made and when it is scheduled to occur. Their study uses data from various clinic types and concludes that shorter scheduling intervals significantly reduce missed appointment rates; authors in [82] reach the same conclusion. Similar conclusions are also reached by the authors in [59], who argue that the the rate of failed intake appointments increases linearly with each day of appointment delay. Interestingly, the results in [100] suggest that increasing waiting times can sometimes be beneficial for reducing no-shows, contrary to previous findings. Here, we consider similar behavior in a maintenance optimization setting.

More specifically, we consider policy-dependent unpunctuality (without the possibility of no-shows). That is, we generalize the model in [64] by allowing the unpunctual behavior be non-stationary. We do so by letting the deviation between the scheduled replacement time and the time when replacement is actually performed depend on the scheduled replacement age, T. We capture this dependence through a non-negative function $z(T)$ that scales the deviation Y multiplicatively so that $Y \cdot z(T)$ is the random deviation (positive or negative) between the prescribed replacement time and the time at which replacement is actually performed.

In our numerical and analytical results, we consider both the case of increasing ($z'(T) > 0$) and decreasing ($z'(T) < 0$) degree of unpunctuality as replacement is scheduled further into the future. An increasing deviation could reflect the worker's forgetfulness as time passes, whereas a decreasing deviation could reflect the worker's ability to better prepare for actions further in the future. Our aim is to compare the optimal long-run cost-rate and the optimal replacement age under non-stationary unpunctual implementation with minimal repair, to those for the cases when the unpunctual behavior is assumed to be either absent or independent of the prescribed replacement time.

Random replacement problems are also closely related to this problem setting [13, 36, 97, 140]. In random replacement problems, replacement is performed at random times as it is here, but it is assumed that a variable work cycle is what prompts replacement times to be random. In contrast, we assume that the potential *unpunctuality* of the maintenance worker causes the actual replacement times to deviate from the intended times in a stochastic way. That said, both problems seek to optimize over a set of replacement time distributions. The random replacement literature assumes a functional form for the distribution that governs the time between replacements, and aims to determine the optimal parameter values for that distribution. We, however, define a set of possible distributions through the random variable Y and function $z(T)$ and seek an optimal replacement time. This difference in defining the set of possible distributions yields fundamentally different models and results.

Herein, we adopt a similar stylized approach to model an age replacement setting in which the actual implementation time of the replacement policy is random and depends on the prescribed policy itself. Studies based on stylized models are typical in the age-replacement maintenance literature; see examples in [57, 136, 141]. Our model enables us to examine how important it is to accurately capture the policy-dependent nature of the maintenance actions. Moreover, we explore how policy-dependent unpunctual implementation affects the long-run cost-rate and the optimal replacement policy. Our analytical results provide insights on how to adjust maintenance policies in anticipation of such unpunctual behavior.

The remainder of the chapter is organized as follows. Section 2.2 provides a general model formulation and establishes [64] and [13, p. 96] as special cases of our model. Section 2.3 provides sufficient conditions for the existence of a unique optimal solution. Section 2.4 makes certain assumptions on the distribution of unpunctuality under stationary and non-stationary behavior and provides analytical results that compare the optimal long-run cost-rate and the optimal prescribed replacement time for our problem with those of stationary unpunctual and punctual implementation. Section 2.5 examines heuristic policies, and provides lower and upper bounds on cost-rate ratios to examine the effect of non-stationary behavior on the long-run cost-rate relative to both stationary unpunctual and punctual implementation. Section 2.6 relaxes the assumptions made in Section 2.4 and analyzes a numerical example that compares the optimal policies and cost-rates of the three problems. Finally, we summarize our findings and discuss future research directions in Section 2.7. The proofs for all results are in Appendix A.

2.2 Model Formulation

Consider a failure-prone system with self-announcing failures that require immediate (instantaneous) minimal repair. Let the continuous random variable X be the time to failure of the system, with c.d.f. $F_X(t)$, p.d.f. $f_X(t)$, mean μ_X, survival function $\bar{F}_X(t)$ and hazard rate function $h_X(t)$, i.e., $h_X(t) = \frac{f_X(t)}{\bar{F}_X(t)}$. We impose the following conditions on $h_X(t)$:

Assumption 1. $h_X(0) = 0$ *and* $h_X(t)$ *is strictly increasing to* $+\infty$*.*

The first part of Assumption 1, $h_X(0) = 0$, implies that the hazard rate is 0 at the time of renewal; this assumption holds for many common failure distributions such as the Weibull distribution. The second condition implies that the hazard rate is strictly increasing and defined for all values of $x \geq 0$. Thus, the random variable X has an infinite support. Both assumptions are commonly used in the maintenance and reliability literature [33, 63, 95, 137, 139].

Next, consider a maintenance planner who prescribes a maintenance policy consisting of an age replacement time, T. That is, preventive replacement is scheduled to be performed once the system has been operating for T units of time; again, failures that occur before age T are repaired minimally. After preventive replacement the system is as-good-as-new. Finally, consider a maintenance worker, i.e., an individual tasked with executing the maintenance policy. We consider the case in which this maintenance worker is unpunctual and the degree of his/her unpunctuality depends on the replacement time T prescribed by the maintenance planner. Note that, we use the terms "non-stationary" and "time-dependent" interchangeably to describe this unpunctual behavior.

Let the probabilistic, non-stationary deviation between the prescribed replacement time and the time at which preventive replacement is actually performed be denoted by the continuous random variable W. To facilitate comparisons to the unpunctual, *stationary* (i.e., *time-independent*) model studied in [64], we use the continuous random variable Y with known c.d.f. $F_Y(y)$ and p.d.f. $f_Y(y)$, and a continuous non-negative function $z(T)$ such that

$$W(T) = Y \cdot z(T). \tag{2.1}$$

Correspondingly, the actual replacement time occurs at age $T + W(T) = T + Y \cdot z(T)$. If $Y > 0$, then replacement is performed later than scheduled, and vice versa if $Y < 0$. We assume that Y is independent of T and capture W's dependence on T via $z(T)$. We impose the following assumptions on Y and $z(T)$:

Assumption 2. *Y is independent of X and is defined over the interval $[a, b]$, $-\infty < a \leq b < \infty$. For all $T \geq 0$, $z(T) \geq 0$ and $z(T)$ is twice differentiable.*

Independence between the random variables X and Y implies that the time to failure distribution does not influence the unpunctual behavior of the maintenance worker. We assume finite bounds on Y for mathematical convenience and that $z(T)$ is non-negative in order to limit the characterization of earliness or lateness to the sign of Y and let $z(T)$

adjust the magnitude of the deviation. Lastly, we assume that $z(T)$ is a smooth, twice differentiable function, which facilitates establishing analytical results. Functional forms of $z(T)$ used for the illustrative examples in this chapter include: $(T/c)^d$, $cT - d + 1$, $\exp(-cT + d)$ for $c > 0$ and $d > 0$, and $c/(T + d)$.

Figure 1 depicts the possible renewal cycle dynamics. In case (i) $z(T)$ is decreasing, i.e., the further into the future that replacement is scheduled, the more punctual the maintenance worker becomes. Conversely, in case (ii) where $z(T)$ is increasing, the further into the future that replacement is scheduled, the less punctual the maintenance worker becomes. In case (iii) the dependence between W and T is absent (i.e., $z(T) = 1$ for all T), which corresponds to the cycle dynamics in [64].

Note that by equation (2.1), if $z(T)$ is increasing (decreasing) over some interval $T \in [T_1, T_2]$, then the absolute value of the mean (if not equal to 0) and variance of $W(T)$ are also increasing (decreasing) over $T \in [T_1, T_2]$. That is, e.g., if Y is defined over the interval $[a, b]$ such that $a < b < 0$, and $z(T)$ is increasing (decreasing) over all T, then the mean of $W(T)$ is decreasing (increasing) and the variance of $W(T)$ is increasing (decreasing) over T. Conversely, e.g., if Y is defined over the interval $[a, b]$ such that $0 < a < b$, and $z(T)$ is increasing (decreasing) over all T, then the mean and variance of $W(T)$ are also increasing (decreasing).

In practice, statistical methods can be used to characterize the random variable Y and function $z(T)$. A possible procedure is as follows. First, a function $z(T)$ is assumed in accordance with historical observations of how the unpunctual behavior changes over time (e.g., monotone increasing or decreasing, non-monotone, linear, concave, or convex). Then, one or multiple classes of distributions is assumed for Y. The parameters of the assumed distribution for Y are estimated from observed data, and the distribution with the maximum likelihood function is adopted. If the fit is poor, this procedure is repeated with new assumptions on $z(T)$ and the classes of distribution for Y [128]. In Section 2.5, we provide bounds on the cost-rate that do not require the full knowledge of the distribution of Y. Hence, these bounds can reflect the value of estimating the true distribution of the unpunctual behavior.

The objective of the maintenance planner is to identify a cost-rate minimizing policy that anticipates the (non-stationary) unpunctual behavior of the maintenance worker. Because replacements return the system to as-good-as-new, we take a renewal-reward approach [111,

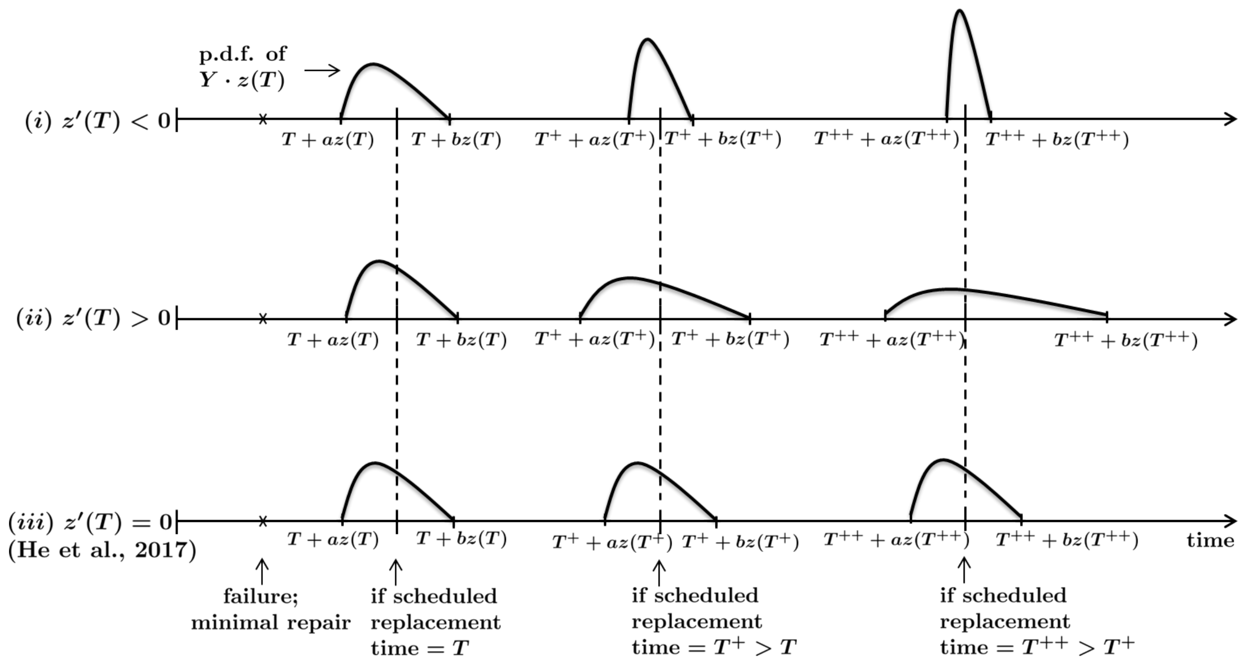


Figure 1: Possible cycle dynamics under age replacement with minimal repair for $a < 0 < b$. The three cases are as follows: (i) $z'(T) < 0$, i.e., the degree of potential unpunctuality of the maintenance worker decreases in the scheduled replacement time; (ii) $z'(T) > 0$, i.e., the degree of potential unpunctuality of the maintenance worker increases in the scheduled replacement time; (iii) $z'(T) = 0$, i.e., the degree of potential unpunctuality of the maintenance worker is stationary. For each case, we depict the distribution of the deviation $T + W(T) = T + Y \cdot z(T)$ for three different ages $T < T^+ < T^{++}$.

p. 52] and formulate the long-run cost-rate as the ratio of the expected renewal cycle cost to the expected renewal cycle length.

Let $C(T)$ and $L(T)$ be the cycle cost and cycle length, respectively. Then, the corresponding long-run average cost-rate is

$$\Omega(T) \equiv \frac{\mathrm{E}_{X,Y}[C(T)]}{\mathrm{E}_{X,Y}[L(T)]}.$$

Because Y can assume negative values, we define the feasible set for T as

$$S = \begin{cases} \{t : t > 0\}, & \text{if } a \geq 0 \\ \{t : t + az(t) > 0\}, & \text{if } a < 0. \end{cases}$$

The decision-making problem for the maintenance planner is given by

$$\min_{T \in S} \Omega(T) = \frac{\mathrm{E}_{X,Y}[C(T)]}{\mathrm{E}_{X,Y}[L(T)]}. \tag{2.2}$$

Similarly, let $\hat{C}(T)$ and $\hat{L}(T)$ be the cycle cost and cycle length, respectively, when the unpunctual behavior is time-independent (i.e., $Y \neq 0$ and $z(T) = 1$ for all T) and let $\hat{\Omega}(T)$ be the corresponding long-run cost-rate. For this case the feasible set for T is

$$\hat{S} = \begin{cases} \{t : t > 0\}, & \text{if } a \geq 0 \\ \{t : t + a > 0\}, & \text{if } a < 0 \end{cases}$$

and the optimization problem is given by

$$\min_{T \in \hat{S}} \hat{\Omega}(T) = \frac{\mathrm{E}_{X,Y}[\hat{C}(T)]}{\mathrm{E}_{X,Y}[\hat{L}(T)]}, \tag{2.3}$$

which is studied in [64].

Lastly, let $\widetilde{C}(T)$ and $\widetilde{L}(T)$ be the cycle cost and cycle length under punctual behavior (i.e., $Y \equiv 0$) and let $\widetilde{\Omega}(T)$ be the corresponding long-run cost-rate. In this case the optimization problem is given by

$$\min_{T > 0} \widetilde{\Omega}(T) = \frac{\mathrm{E}_X[\widetilde{C}(T)]}{\mathrm{E}_X[\widetilde{L}(T)]}, \tag{2.4}$$

a well studied classical model (see [13]).

Let c_p and c_m denote the replacement cost and the minimal repair cost, respectively. When replacement is always performed on time,

$$\mathrm{E}_X[\widetilde{C}(T)] = c_m \int_0^T h_X(x)dx + c_p \ , \quad \mathrm{E}_X[\widetilde{L}(T)] = T,$$

and the optimal solution to problem (2.4), $\widetilde{T}^*$, corresponds to solving

$$\min_{T>0} \widetilde{\Omega}(T) = \frac{c_m \int_0^T h_X(x)dx + c_p}{T}. \tag{2.5}$$

If the timing of replacement is unpunctual, but independent of the scheduled replacement time, renewal occurs at time $T + Y$. Therefore,

$$\begin{aligned}
\mathrm{E}_{X,Y}[\hat{C}(T)] &= \int_a^b \Big(c_m \int_0^{T+y} h_X(x)dx\Big) dF_Y(y) + c_p, \\
\mathrm{E}_{X,Y}[\hat{L}(T)] &= T + \mu_Y,
\end{aligned}$$

and the optimal solution to problem (2.3), $\hat{T}^*$, corresponds to solving

$$\min_{T\in\hat{S}} \hat{\Omega}(T) = \frac{\int_a^b \Big(c_m \int_0^{T+y} h_X(x)dx\Big) dF_Y(y) + c_p}{T + \mu_Y}. \tag{2.6}$$

When the degree to which replacement is performed early or late depends on the scheduled replacement time, renewal occurs at time $T + Y \cdot z(T)$ and the expected cycle length and expected cycle cost are given by

$$\begin{aligned}
\mathrm{E}_{X,Y}[C(T)] &= \int_a^b \Big(c_m \int_0^{T+yz(T)} h_X(x)dx\Big) dF_Y(y) + c_p, \\
\mathrm{E}_{X,Y}[L(T)] &= T + \mu_Y z(T),
\end{aligned}$$

respectively, and the optimal solution to problem (2.2), T^*, corresponds to optimizing

$$\min_{T\in S} \Omega(T) = \frac{\int_a^b \Big(c_m \int_0^{T+yz(T)} h_X(x)dx\Big) dF_Y(y) + c_p}{T + \mu_Y z(T)}. \tag{2.7}$$

Note that the terms

$$\int_a^b \Big(c_m \int_0^{T+yz(T)} h_X(x)dx\Big) dF_Y(y) / \big(T + \mu_Y z(T)\big)$$

and $c_p/\big(T + \mu_Y z(T)\big)$ represent the long-run minimal repair cost-rate and long-run replacement cost-rate, respectively.

Table 1 summarizes the notation for the three models presented in Section 2.2.

Table 1: Summary of age replacement models with minimal repair under punctual replacement [13]; stationary, unpunctual replacement [64]; and non-stationary, unpunctual replacement (our model).

	Model in [13]	Model in [64]	Our model
Cost-rate	$\widetilde{\Omega}(T)$	$\hat{\Omega}(T)$	$\Omega(T)$
Expected cycle cost	$c_m \int_0^T h_X(x)dx + c_p$	$\int_a^b \left(c_m \int_0^{T+y} h_X(x)dx \right) dF_Y(y) + c_p$	$\int_a^b \left(c_m \int_0^{T+yz(T)} h_X(x)dx \right) dF_Y(y) + c_p$
Expected cycle length	T	$T + \mu_Y$	$T + \mu_Y z(T)$
Prescribed replacement time	$\widetilde{T}$	$\hat{T}$	T
Actual replacement time	$\widetilde{T}$	$\hat{T} + Y$	$T + Y \cdot z(T)$

2.3 Existence of a Unique Optimal Solution

In this section, we establish conditions under which problem (2.7) has a unique solution. Setting the first derivative of the objective function in problem (2.7) equal to zero yields

$$m(T) = k, \tag{2.8}$$

where $k = \frac{c_p}{c_m} > 1$ and

$$\begin{aligned} m(T) = &- \int_a^b \int_0^{T+yz(T)} h_X(x) f_Y(y) dx dy \\ &+ \frac{T + \mu_Y z(T)}{1 + \mu_Y z'(T)} \int_a^b (1 + yz'(T)) h_X(T + yz(T)) f_Y(y) dy. \end{aligned}$$

If $m(T) = k$ has a unique solution, then that solution, T^*, satisfies equation (2.8) and the minimum cost-rate $\Omega(T^*)$ is

$$\frac{c_m}{1 + \mu_Y z'(T^*)} \int_a^b (1 + yz'(T^*)) h_X(T^* + yz(T^*)) f_Y(y) dy.$$

Proposition 1 establishes sufficient conditions to ensure that there exists a unique optimal solution for problem (2.7).

Proposition 1. *For all $T \in S$, let (i) $1 + \mu_Y z'(T) \geq \delta > 0$ for some $\delta \in \mathbb{R}$, and (ii) $z''(T) \geq 0$. If*

$$\lim_{T \to \max\{\min\{t:\, t+az(t)>0\},0\}} m(T) < k, \tag{2.9}$$

then $\Omega(T)$ is quasi-convex and there exists a unique solution T^* to (2.7). Otherwise,

$$\inf \Omega(T) = \lim_{T \to \max\{\min\{t:\, t+az(t)>0\},0\}} \Omega(T). \tag{2.10}$$

Recall that the proofs of all analytical results including Proposition 1 are provided in Appendix A. The sufficient condition (i) in Proposition 1 requires $1 + \mu_Y z'(T)$ to be strictly positive. Note that this expression is the derivative of the expected replacement time, $T + \mu_Y z(T)$. Hence, the condition implies that the later the prescribed replacement time, the later the expected replacement time, which is a reasonable assumption.

The sufficient condition (ii) in Proposition 1 requires $z(T)$ to be convex. Although, we are unable to prove the uniqueness of an optimal solution without this condition, all of our numerical instances suggest that relaxing this condition does not affect the uniqueness of the solution to problem (2.7). Violation of $1 + \mu_Y z'(T) > 0$, on the other hand, can result in multiple solutions to equation (2.8); see Example 1.

Example 1. Let $X \sim$ Weibull$(4, 1)$, $Y \sim$ Uniform$[-10, -9]$, $z(T) = \left(\frac{T}{10.54}\right)^2$ and $k = 4$. As seen in Figure 2, there are two solutions to equation (2.8), $T = T^* = 1.20$ and $T = 10.51$. The third extremum of $\Omega(T)$ occurs at age 5.87 where $1 + \mu_Y z'(5.87) = 0$.

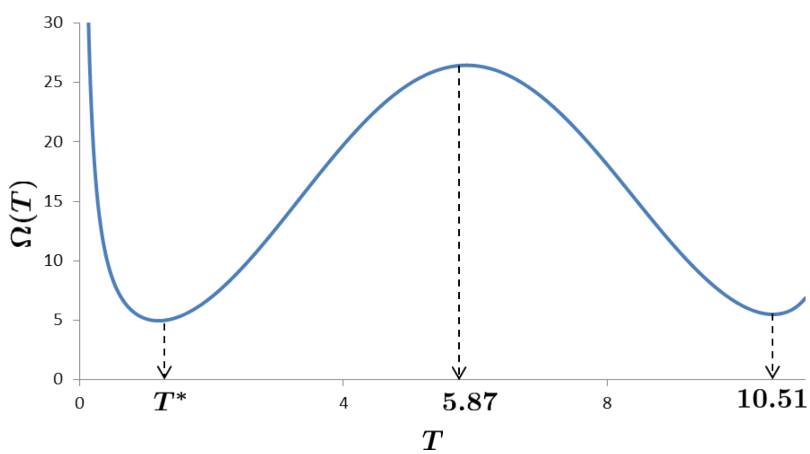


Figure 2: Expected long-run cost-rate for Example 1. Because $1 + \mu_Y z'(T) > 0$ does not hold for all $T \in S$, there are two local minima for $\Omega(T)$: $T = T^* = 1.20$ and $T = 10.51$.

Under the parameter settings in Example 1, $1+\mu_Y z'(T) > 0$ holds for values of $T < 5.87$ and hence, the expected replacement time (i.e., $T + \mu_Y z(T)$) is increasing for $T < 5.87$. Conversely, $1+\mu_Y z'(T) < 0$ holds for values of $T > 5.87$ and hence, the expected replacement time is decreasing for $T > 5.87$; i.e., the later the prescribed replacement time, the earlier the expected replacement time. Because condition (i) in Proposition 1 is violated, there are two local minima for problem (2.7); one at $T = 1.20$ where the expected replacement time is increasing, and another at $T = 10.51$ where the expected replacement time is decreasing. The local minimum at $T = 1.20$ corresponds to the global minimum to problem (2.7). ■

Based on Proposition 1, we impose the following condition:

Assumption 3. *There exists a* $\delta \in \mathbb{R}$ *such that* $1 + \mu_Y z'(T) \geq \delta > 0$ *for all* $T \in S$*.*

In all of the analytical results and numerical examples to follow, we ensure that Assumption 3 and condition (2.9) hold. Note that the value of k must be relatively large for condition (2.9) to hold. That is, condition (2.9) may not hold when the preventive replacement cost c_p is nearly equal to the minimal repair cost c_m; however, $c_p \approx c_m$ is unlikely to hold in practice.

2.4 Analytical Results

In this section, we derive analytical results to compare the optimal cost-rates and optimal solutions of problems (2.5), (2.6) and (2.7). These comparisons inform maintenance planners as to how to adjust their policies in anticipation of unpunctual maintenance workers. In general, it is not straightforward to compare the optimal long-run cost-rates and replacement policies for two arbitrary instances of problems (2.6) and (2.7) analytically (see the numerical example in Section 2.6). Hence, throughout Section 2.4, to compare the optimal long-run cost-rate and replacement policies of problems (2.6) and (2.7), we often assume that either

$$z(\hat{T}^*) = 1, \tag{2.11}$$

or

$$z(T^*) = 1. \tag{2.12}$$

Based on equation (2.1), condition (2.11) implies that at time $\hat{T}^*$ the degree of unpunctuality for problems (2.6) and (2.7) are stochastically equivalent and hence $\Omega(\hat{T}^*) = \hat{\Omega}(\hat{T}^*)$. That is, although in the non-stationary case the deviation distribution changes over time, it coincides with that of the stationary case at $\hat{T}^*$. More intuitively, if the worker exhibits *non-stationary* unpunctual behavior and the system is intended to be replaced at the optimal replacement time under *stationary* unpunctual behavior (i.e., $\hat{T}^*$), then the cost-rate would match that of a worker who exhibits stationary unpunctual behavior implementing the same replacement policy; for scheduled replacement times before or after

$\hat{T}^*$, however, the deviation between planned and actual time may be smaller or larger than that of a maintenance worker who exhibits stationary unpunctual behavior. The condition in equation (2.12) can be explained similarly. In Section 2.6, we provide a numerical example to examine the case in which neither (2.11) nor (2.12) holds.

2.4.1 Analysis of $\boldsymbol{\Omega(T^*)}$ vs. $\boldsymbol{\hat{\Omega}(\hat{T}^*)}$ vs. $\boldsymbol{\widetilde{\Omega}(\widetilde{T}^*)}$

Barlow and Proschan (1965) establish the intuitive fact that the long-run cost-rate under uncertain timing of replacements is greater than it would be under deterministically timed replacements. Theorem 1 states this fact in the context of our problem in which replacement times are uncertain with a time-dependent distribution.

Theorem 1. $\widetilde{\Omega}(\widetilde{T}^*) \leq \Omega(T^*)$.

Theorem 2 compares the optimal long-run cost-rates of problems (2.5), (2.6) and (2.7).

Theorem 2. *If* $z(\hat{T}^*) = 1$, *then* $\widetilde{\Omega}(\widetilde{T}^*) \leq \Omega(T^*) \leq \hat{\Omega}(\hat{T}^*)$.

Similar to Theorem 1, the first inequality in Theorem 2 intuitively establishes that when the timing of replacement is certain, the optimal long-run cost-rate is smaller than it would be under unpunctual replacement. Therefore, the cost-rate of problem (2.5) is less than that of both problems (2.6) and (2.7). The second inequality establishes that with $z(\hat{T}^*) = 1$, the optimal cost-rate under non-stationary unpunctual behavior is no more than that under stationary unpunctual behavior. Theorem 3 compares the optimal solutions under the assumption that $z(T^*) = 1$ instead.

Theorem 3. *If* $z(T^*) = 1$, *then* $\widetilde{\Omega}(\widetilde{T}^*) \leq \hat{\Omega}(\hat{T}^*) \leq \Omega(T^*)$.

In Theorem 3, with $z(T^*) = 1$, the cost-rates of problems (2.6) and (2.7) coincide at the optimal solution to the problem with non-stationary unpunctual behavior. Hence, the optimal solution to the problem with stationary unpunctual behavior is at least as good as that for the non-stationary case. Intuitively, the fact that f_Y is stationary gives the maintenance planner more flexibility in obtaining the optimal cycle length and cost as opposed to $f_{W(T)}$, and results in a lower cost per unit time.

The insights gained from Theorems 2 and 3 are two-fold. First, the optimal cost-rate under stationary unpunctual behavior can be improved by a maintenance worker with non-stationary unpunctual behavior if the mean or variance of her/his unpunctuality at the optimal replacement time is no more than that of the worker with stationary behavior. This result holds because by exploiting the non-stationary behavior, the maintenance worker can decrease the mean and variance of unpunctuality by scheduling replacement earlier or later than $\hat{T}^*$. Second, the optimal cost-rate under non-stationary unpunctual behavior can be improved by a maintenance worker with stationary unpunctual behavior if the mean or variance of her/his unpunctuality is no more than that of the worker with non-stationary behavior at the optimal replacement time. This result holds because with a more consistently behaved worker, the maintenance planner can obtain a better cycle length.

Next, under some assumptions on the distribution of Y, we compare the cost-rate functions of problems (2.6) and (2.7) over all T, not just at their optimal solutions as in Theorems 2 and 3.

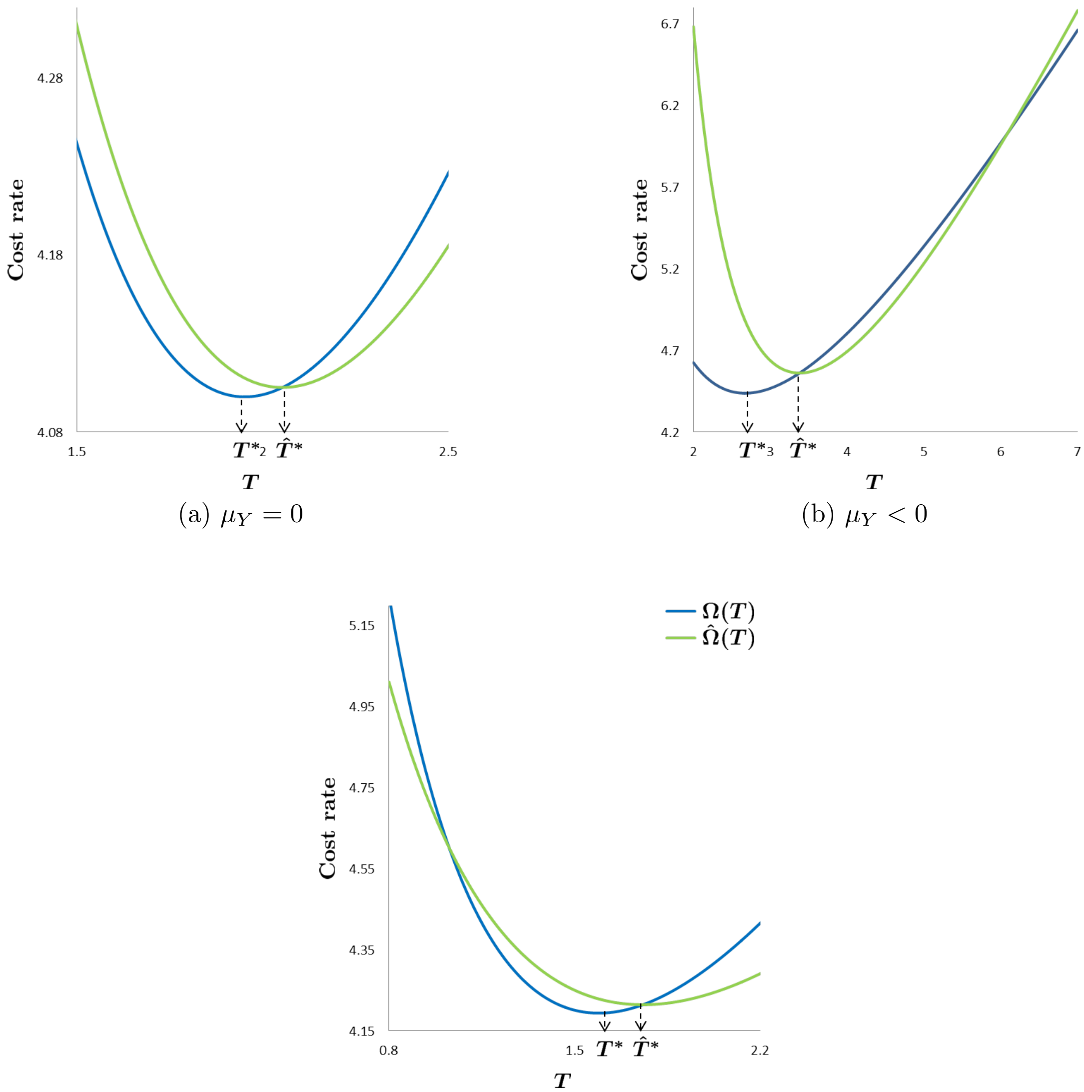
Theorem 4. *Let $\mu_Y(y) = 0$ and $f_Y(y)$ be symmetric. For all T such that $z(T) \leq 1$, $\Omega(T) \leq \hat{\Omega}(T)$. Conversely, for all T such that $z(T) \geq 1$, $\Omega(T) \geq \hat{\Omega}(T)$.*

For $\mu_Y = 0$ the expected cycle lengths of problem (2.6) and (2.7) are equal. Note that under non-stationary unpunctual behavior, if replacement is prescribed at a time T for which $z(T) \geq 1$, then the variance of the degree of unpunctuality is at least as large as that for the stationary case. Theorem 4 states that under such unpunctual behavior, larger variance in the degree of unpunctuality results in a larger expected cycle cost. The opposite holds if replacement is prescribed at a time T with $z(T) \leq 1$.

Example 2 compares $\Omega(T)$ and $\hat{\Omega}(T)$ for (*i*) $\mu_Y = 0$ and symmetric $f_Y(y)$, (*ii*) $\mu_Y < 0$ and (*iii*) $\mu_Y > 0$ under a monotone function $z(T)$.

Example 2. Let $z(T) = \frac{T}{\hat{T}^*}$. Assume $k = 4$ and $X \sim$ Weibull$(2, 1)$. Let (*i*) $Y \sim$ Uniform$[-0.8, 0.8]$, (*ii*) $Y \sim$ Uniform$[-3, 0.8]$ and (*iii*) $Y \sim$ Uniform$[-0.8, 1.5]$. For these scenarios, see Figures 3(a), 3(b) and 3(c), respectively.

First, note that because $z(\hat{T}^*) = 1$, as stated in Theorem 2, in all three cases in Figure 3 the optimal cost-rate under non-stationary unpunctuality is less than that under stationary



(a) $\mu_Y = 0$

(b) $\mu_Y < 0$

(c) $\mu_Y > 0$

Figure 3: Long-run cost-rate under non-stationary and stationary unpunctual behavior for Example 2.

unpunctuality. In Figure 3(a), $\mu_Y = 0$ and $f_Y(y)$ is symmetric; thus, the conditions of Theorem 4 hold and because $z(T)$ is monotone, the cost-rate functions coincide only at $\hat{T}^*$. However, for $\mu_Y < 0$ in Figure 3(b) and $\mu_Y > 0$ in Figure 3(c), the conditions of Theorem 4 do not hold and the functions coincide twice. In Figure 3(c), although the variance of the degree of unpunctuality is smaller for all replacement ages T less than $\hat{T}^*$, the expected cycle length is also smaller under non-stationary unpunctual behavior; as a result, for $T < 1.03$, the corresponding cost-rate function is larger than the time-independent unpunctual cost-rate. The opposite holds for Figure 3(b), where $\mu_Y < 0$. ■

2.4.2 Analysis of $\boldsymbol{T^*}$ vs. $\boldsymbol{\hat{T}^*}$ vs. $\boldsymbol{\widetilde{T}^*}$

In this section, we compare the optimal replacement ages for problems (2.5), (2.6) and (2.7). First, we establish conditions under which we can analytically compare the optimal replacement ages under stationary (problem (2.6)) and non-stationary (problem (2.7)) unpunctuality. Similar to Theorems 2 and 3, we assume that either $z(\hat{T}^*) = 1$ or $z(T^*) = 1$, respectively, so that we are able to compare the corresponding optimal solutions. Note that throughout this section, the convexity of $z(T)$ is a sufficient condition for the analytical results to hold, but all of our numerical instances suggest that relaxing this condition does not affect the results.

Theorem 5. *Let $z(\hat{T}^*) = 1$ and $z''(T) \geq 0$ for all $T \in S$. If $z'(\hat{T}^*) > 0$, then $T^* < \hat{T}^*$. Conversely, if $z'(\hat{T}^*) \leq 0$, then $T^* \geq \hat{T}^*$.*

The intuition behind Theorem 5 is similar to that behind Theorem 2. For $z(T)$ increasing, if the maintenance planner schedules replacement earlier than $\hat{T}^*$, then the mean and variance of the degree of unpunctuality decrease, resulting in a lower cost-rate. The opposite holds for decreasing $z(T)$; Corollary 1 establishes the analogous result for $z(T^*) = 1$.

Corollary 1. *Let $z(T^*) = 1$ and $z''(T) \geq 0$ for all $T \in S$. If $z'(T^*) > 0$, then $T^* < \hat{T}^*$. Conversely, if $z'(T^*) \leq 0$, then $T^* \geq \hat{T}^*$.*

Next, Example 3 demonstrates the relationship between the cost-rate functions and optimal solutions for problems (2.6) and (2.7) when either $z(\hat{T}_1^*) = 1$ or $z(T^*) = 1$, where $\hat{T}_1^*$ is the solution that minimizes a stationary unpunctual cost-rate $\hat{\Omega}_1(T)$. Under the condition that $z(\hat{T}_1^*) = 1$, we solve for the optimal solution T^* that minimizes the non-stationary unpunctual cost-rate $\Omega(T)$. Lastly, we solve for the optimal solution $\hat{T}_2^*$ that minimizes a second stationary unpunctual cost-rate $\hat{\Omega}_2(T)$ under the condition that the distribution of unpunctual behavior is equal to $f_{W(T^*)}$ (i.e., $z(T^*) = 1$).

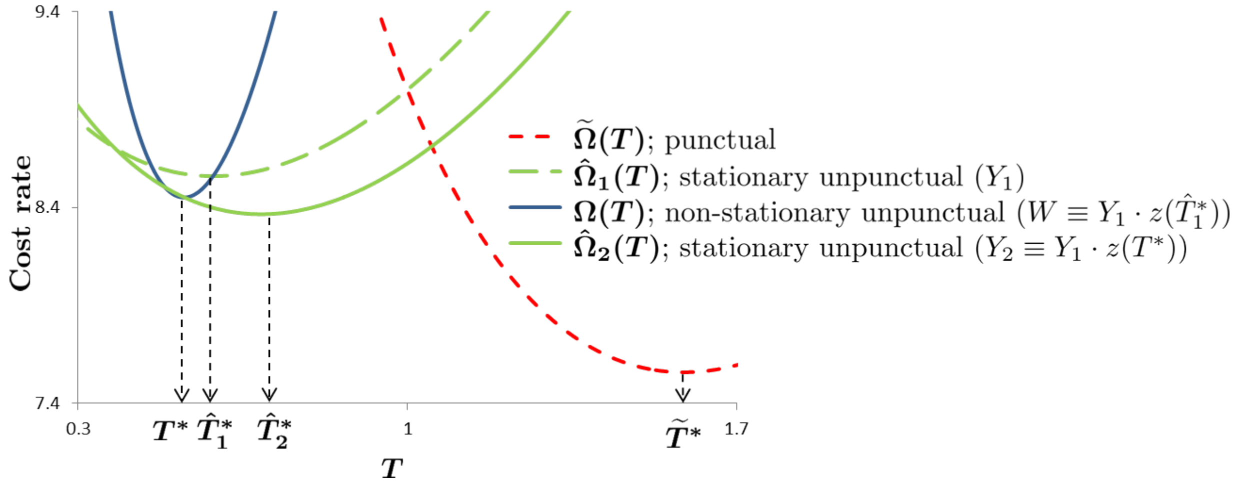


Figure 4: Long-run cost-rate functions under punctual behavior $\big($i.e., $\widetilde{\Omega}(T)\big)$, non-stationary unpunctual behavior $\big($i.e., $\Omega(T)\big)$, stationary unpunctual behavior for $Y_1 \sim$ Uniform$[0,2]$ $\big($i.e., $\hat{\Omega}_1(T)\big)$ and stationary unpunctual behavior for $Y_2 \equiv z(T^*) \cdot Y_1$ $\big($i.e., $\hat{\Omega}_2(T)\big)$ for the problem setting described in Example 3. Note that when comparing $\Omega(T)$ and $\hat{\Omega}_1(T)$, $z(\hat{T}_1^*) = 1$; however, when comparing $\Omega(T)$ and $\hat{\Omega}_2(T)$, $z(\hat{T}_2^*) \neq 1$.

Example 3. Let $X \sim$ Weibull$(3,1)$ and $k = 8$. Then, we have $\widetilde{\Omega}(\widetilde{T}^*) = 7.56$ and $\widetilde{T}^* = 1.59$. First, we let $Y_1 \sim$ Uniform$[0,2]$ and obtain the optimal cost-rate and replacement time under stationary unpunctual behavior: $\hat{\Omega}_1(\hat{T}^*) = 8.56$, $\hat{T}_1^* = 0.59$. Next, we let $z(T) = \frac{T}{\hat{T}_1^*}$, and $\Omega(T^*) = 8.45$ and $T^* = 0.53$. Finally, we set $Y_2 \equiv W(T^*) \equiv z(T^*) \cdot Y_1$; that is, we let the distribution of the stationary unpunctual behavior be identical to the distribution of the non-stationary unpunctual behavior at the optimal replacement time T^*, for which $\hat{\Omega}_2^* = 7.82$ and $\hat{T}_2^* = 0.64$. Figure 4 depicts the long-run cost-rate function for each problem instance.

In Figure 4, when comparing $\Omega(T)$ to $\hat{\Omega}_1(T)$, the optimal replacement time under non-stationary behavior is earlier than that under stationary behavior. This relationship holds because $z(T) = \frac{T}{\hat{T}_1^*}$ and by replacing earlier than $\hat{T}_1^*$, the mean and variance of the degree of deviation would be smaller (recall Theorem 5). Furthermore, the optimal cost-rate under non-stationary unpunctual behavior is smaller than that under stationary unpunctual behavior because the mean and variance of the degree of unpunctuality under non-stationary behavior at T^* are smaller than that of stationary behavior (recall Theorem 2).

In contrast, when comparing $\Omega(T)$ to $\hat{\Omega}_2(T)$, the optimal cost-rate under non-stationary behavior is larger than its stationary counterpart. This relationship holds because $z(\hat{T}_2^*) \neq 1$ and the deviation under stationary behavior has the same distribution as that under non-stationary behavior at the optimal replacement age T^* (recall Theorem 3). Moreover, $\hat{T}_2^*$ is closer to $\widetilde{T}^*$ compared to $\hat{T}_1^*$. This intuitive result holds because the mean and variance of Y_2 for $\hat{\Omega}_2(T)$ is smaller than that of Y_1 for $\hat{\Omega}_1(T)$; hence, the optimal solution that minimizes $\hat{\Omega}_2(T)$ is nearer to that under the case of punctual implementation. ■

Next, to analyze the relative magnitude of T^*, $\hat{T}^*$ and $\widetilde{T}^*$, we exploit the following four results from [64], stated here as Lemmas 1-4.

Lemma 1 [64]. *If $\mu_Y = 0$ and $h_X(t)$ is concave, then $\hat{T}^* \geq \widetilde{T}^*$.*

Lemma 2 [64]. *If $a < b \leq 0$, $h_X(t)$ is convex and (i) $\widetilde{m}(-a) \geq k$, then $\widetilde{T}^* \leq -a < T^*$; (ii) $\widetilde{m}(-a) < k$, then $-a < \widetilde{T}^* < \hat{T}^*$.*

Lemma 3 [64]. *If $0 \leq a < b$, $h_X(t)$ is convex and $\lim_{T\to+0} \hat{m}(T) \geq 0$, then $0 < \hat{T}^* < \widetilde{T}^*$.*

Lemma 4 [64]. *If $X \sim$ Weibull(α, β), $\alpha \in \mathbb{Z}_+$, $\alpha > 2$ and $f_Y(y)$ is symmetric w.r.t. $y = 0$, then $\hat{T}^* \leq \widetilde{T}^* = (\frac{k}{\alpha-1})^{\frac{1}{\alpha}}\beta$. Moreover, if $\alpha = 3$, then $\hat{T}^* = \widetilde{T}^* = (\frac{k}{2})^{\frac{1}{3}}\beta$.*

The function $\widetilde{m}(T)$ in Lemma 2 is obtained by setting the first derivative of the objective function in problem (2.5) equal to zero:

$$\widetilde{m}(T) \equiv T \cdot h_X(T) - \int_0^T h_X(x)dx = k. \tag{2.13}$$

The solution to equation (2.13) is $\widetilde{T}^*$. Similarly, the function $\hat{m}(T)$ in Lemma 3 is obtained by setting the first derivative of the objective function in problem (2.6) equal to zero:

$$\begin{aligned}\hat{m}(T) \equiv (T + \mu_Y) &\int_a^b h_X(T + y) f_Y(y)dy \\ &- \int_a^b \int_0^{T+y} h_X(x) f_Y(y)dxdy = k.\end{aligned} \tag{2.14}$$

The solution to equation (2.14) is $\hat{T}^*$.

We use Theorem 5 and Lemmas 1-4 to establish the following propositions that compare the optimal solutions of problems (2.5), (2.6) and (2.7). For Propositions 2-5 we assume that $z(\hat{T}^*) = 1$.

Proposition 2. *Let $z''(T) \geq 0$, $\mu_Y = 0$ and $h_X(t)$ be concave. If $z'(\hat{T}^*) < 0$, then $\widetilde{T}^* < \hat{T}^* < T^*$.*

The concavity of $h_X(t)$ implies that the rate of increase in the hazard function decreases over time. Under such a hazard rate, if the expected replacement time is equal to the prescribed replacement time (i.e., $\mu_Y = 0$), then it is optimal to schedule replacement later than we would under punctual behavior (i.e., $\widetilde{T}^* < \hat{T}^*$). Moreover, if the variance of the degree of unpunctuality decreases as replacement is scheduled further into the future, then it is optimal to schedule the replacement later than we would under stationary unpunctual behavior.

Example 4 demonstrates that if instead $z'(\hat{T}^*) > 0$, then depending on the model parameter values, both replacing earlier ($T^* < \widetilde{T}^* < \hat{T}^*$) and replacing later ($\widetilde{T}^* < T^* < \hat{T}^*$) than the optimal replacement age under punctual implementation may be optimal.

Example 4. Let $k = 5$, $X \sim$ Weibull$(1.5, 1)$ and $Y \sim$ Uniform$[-4, 4]$. The optimal solutions for problems (2.5) and (2.6) are $\widetilde{T}^* = 4.64$ and $\hat{T}^* = 5.45$, respectively, and clearly $\widetilde{T}^* < \hat{T}^*$. Consider (i) $z(T) = \left(\frac{T}{\hat{T}^*}\right)^2$ and (ii) $z(T) = 0.4\frac{T}{\hat{T}^*} + 0.6$. In case (i) $T^* = 4.11$ and hence $T^* < \widetilde{T}^*$. However, in case (ii) $T^* = 5.04$ and $T^* > \widetilde{T}^*$. Note that, in case (i), if the maintenance planner disregards the non-stationary behavior and only anticipates stationary unpunctual behavior, then replacement is prescribed at a time later than $\widetilde{T}^*$. However, under a true characterization of $z(T)$, preventive replacement is prescribed earlier than $\widetilde{T}^*$. ■

Proposition 3. *Let $z''(T) \geq 0$, $a < b \leq 0$ and $h_X(t)$ be convex. If $z'(\hat{T}^*) < 0$, then $\widetilde{T}^* < \hat{T}^* < T^*$.*

Proposition 3 establishes conditions under which, if the maintenance worker never performs replacement later than intended, then the maintenance planner should schedule replacement later than he would under punctual behavior (i.e., $\widetilde{T}^* < \hat{T}^*$). Moreover, if the mean and variance of the degree of unpunctuality decreases when replacement is scheduled beyond $\hat{T}^*$ (i.e., $z'(\hat{T}^*) < 0$), then it is optimal to shift replacement even later than that under stationary unpunctual behavior (i.e., $\hat{T}^* < T^*$). Conversely, Proposition 4 implies that if the maintenance worker never performs replacement earlier than intended, then the maintenance planner should schedule replacement earlier than he would under punctual behavior.

Proposition 4. *Let $z''(T) \geq 0$, $0 \leq a < b$, $h_X(t)$ be convex and $\lim_{T \to +0} \hat{m}(T) \geq 0$. If $z'(\hat{T}^*) > 0$, then $T^* < \hat{T}^* < \widetilde{T}^*$.*

Proposition 4 addresses the opposite scenario of Proposition 3, however, unlike Proposition 3, the result in Proposition 4 depends on the distribution of the delay time Y and requires $\lim_{T \to +0} \hat{m}(T) \geq 0$. In [64], it is shown that this condition holds for sufficiently large values of the mean delay time μ_Y. For some instances with large variation in delay but a small mean delay time, the majority of the cost-rate is attributable to minimal repairs. Therefore, the long-run cost-rate decreases when replacement is prescribed at an age greater than $\widetilde{T}^*$. Hence, it may be optimal to schedule the replacement later than we would under punctual behavior, even though the worker is never early (see Example 1 in [64]).

In Propositions 3 and 4, the comparison between the optimal replacement times of problems (2.6) and (2.7) is a direct result of Theorem 5 for the given conditions on $z'(\hat{T}^*)$. If either $z'(\hat{T}^*) > 0$ in Proposition 3 or $z'(\hat{T}^*) < 0$ in Proposition 4 is violated, however, then the relationships between the optimal replacement times under non-stationary unpunctuality and punctual implementation still hold for all of the numerical instances that we tested.

Proposition 5. *Let $z''(T) \geq 0$, $X \sim$ Weibull(α, β), $\alpha \in \mathbb{Z}_+$, $\alpha > 2$ and $f_Y(y)$ be symmetric w.r.t. $y = 0$. If $z'(\hat{T}^*) > 0$, then $T^* < \hat{T}^* \leq \widetilde{T}^*$. Moreover, for the special case of $\alpha = 3$, (i) if $z'(\hat{T}^*) > 0$, then $T^* < \hat{T}^* = \widetilde{T}^* = (\frac{k}{2})^{\frac{1}{3}}\beta$, and (ii) if $z'(\hat{T}^*) < 0$, then $\hat{T}^* = \widetilde{T}^* < T^*$.*

Note that a Weibull distribution with shape parameter greater than two has a convex hazard function. Proposition 5 shows that if the hazard rate increases sufficiently quickly (i.e., $\alpha > 2$) and the unpunctual behavior of the maintenance worker is constant over time, then it is optimal to prescribe replacement at an earlier age compared to the punctual implementation case. Furthermore, under a symmetric $f_Y(y)$ and $\alpha = 3$, the optimal solutions for stationary unpunctual and punctual implementation coincide [64]. However, the optimal solution under non-stationary implementation is different and can be obtained by solving $T^3 + 3T^2 z(T) z'(T) \sigma_Y = \frac{k}{2}\beta^3$ for T.

In general, if $X \sim$ Weibull(α, β) for $\alpha \neq 3$ and $z'(\hat{T}^*) < 0$, it is difficult to characterize the relationship between T^* and $\widetilde{T}^*$. Example 5 illustrates the type of results than can be obtained for this case under different forms of $z(T)$.

Example 5. Let $X \sim$ Weibull$(5, 2)$, $Y \sim$ Uniform$[-1, 1]$ and $k = 16$. The optimal solutions for problems (2.5) and (2.6) are $\widetilde{T}^* = 2.63$ and $\hat{T}^* = 2.52$. Consider (i) $z(T) = \exp(-2\frac{T}{\hat{T}^*}+2)$ and (ii) $z(T) = \exp(-\frac{1}{2}\frac{T}{\hat{T}^*} + \frac{1}{2})$. In case (i) $T^* = 2.76$ and hence $T^* > \widetilde{T}^*$. However, in case (ii), $T^* = 2.59 < \widetilde{T}^*$. ■

To summarize, our derivations and examples indicate that if the maintenance worker with stationary or non-stationary unpunctual behavior is never late, then it is usually optimal to schedule the replacement time later than we would for a punctual worker. In many instances, the opposite holds for a maintenance worker who is never early; however, in some scenarios where the degree of unpunctuality has a small mean and large variance, it may be optimal to schedule replacement later than we would under the punctual behavior. Finally, if the maintenance worker is sometimes late and sometimes early (degree of unpunctuality varies around zero), then the optimal replacement times under stationary and non-stationary unpunctual behavior can have different relationships with $\widetilde{T}^*$. For these scenarios, ignoring the non-stationary behavior can increase the cost-rate significantly.

2.5 Heuristic Policies and Cost-rate Bounds

In this section, we examine heuristic policies and provide some bounds on cost-rate ratios to study the effect of non-stationary behavior on the cost-rate. Note that the results in this section generally do not require $z(T^*) = 1$ or $z(\hat{T}^*) = 1$. That is the results in this section hold for a general function $z(T)$ with the exception of Proposition 7.

2.5.1 Heuristic Policies

Full characterization of an individual's unpunctual behavior can be difficult. Therefore, heuristic policies that do not depend on a fully specified f_Y can be appealing. Hence next, we numerically study the percent increase in cost-rate induced by prescribing replacement at such a time. First, consider prescribing replacement at age $\widetilde{T}^*$, which ignores the unpunctual behavior. In Section 2.5.2 we quantify the cost-rate ratio for this heuristic (Proposition 8) and provide numerous examples that illustrate the percent increase in cost-rate over the optimal policy.

Next, consider opting to replace at $\widetilde{T}^* - \mu_Y$. This heuristic, which is discussed in [64], only anticipates the expected degree of stationary unpunctuality. In Example 6, we show that the increase in cost-rate under this policy can be as large as 62.97% under certain parameter values.

Example 6. Let $k = 20$, $X \sim$ Weibull$(8, 10)$, $Y \sim$ Uniform$[-3, 6]$ and $z(T) = (\frac{T}{8})^2$. The optimal solutions for problems (2.5) and (2.7) are $\widetilde{T}^* = 11.40$ and $T^* = 7.95$. For this problem instance, replacing at age $\widetilde{T}^* - \mu_Y$ increases the cost-rate by 62.97% (i.e., $\Omega(\widetilde{T}^* - \mu_Y)/\Omega(T^*) = 1.6297$). ■

Finally, consider the heuristics $\widetilde{T}_1 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^* - \mu_Y)$ and $\widetilde{T}_2 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^*)$ which appear to perform well in our numerical experimentation because both policies anticipate the expected degree of unpunctuality under non-stationary behavior. That is, policy $\widetilde{T}_1$ subtracts the expected degree of unpunctuality at time $\widetilde{T}^* - \mu_Y$ from $\widetilde{T}^*$, and policy $\widetilde{T}_2$ subtracts the expected degree of unpunctuality at time $\widetilde{T}^*$ from $\widetilde{T}^*$. In Example 6, opting for the sub-optimal replacement time $\widetilde{T}_1$ yields $\Omega(\widetilde{T}_1)/\Omega(T^*)$=1.1659 and opting for the sub-optimal replacement time $\widetilde{T}_2$ yields $\Omega(\widetilde{T}_2)/\Omega(T^*)$=1.0161. Hence, these heuristic policies appear to not affect the long-run cost-rate significantly.

The superiority of heuristic policy $\widetilde{T}_1$ or $\widetilde{T}_2$ over the other is problem specific. Hence, we execute a designed experiment to generate insights as to when one outperforms the other. For this numerical experiment, let $X \sim$ Weibull$(\alpha, 10)$ with $\alpha \in \{1.5, 2, 4, 8\}$; $k \in \{2, 4, 6, ..., 20\}$; $Y \sim$ Uniform$[a, b]$ with a and b as specified in Figure 5; and $z(T)$ be of the three forms given in Figure 5. Figure 5 presents, for each combination of $z(T)$, $[a, b]$ and α, the heuristic that yields the smaller expected cost-rate. Because there are no instances for which the value of k affects the outcome, we omit k from the figure.

For $z(T) = (\frac{T}{\hat{T}^*})^2$ it is apparent that $\widetilde{T}_1$ usually yields a lower expected cost-rate compared to $\widetilde{T}_2$. For $z(T) = \sqrt{\frac{T}{\hat{T}^*}}$ and small values of α, $\widetilde{T}_1$ dominates $\widetilde{T}_2$; the opposite holds for larger

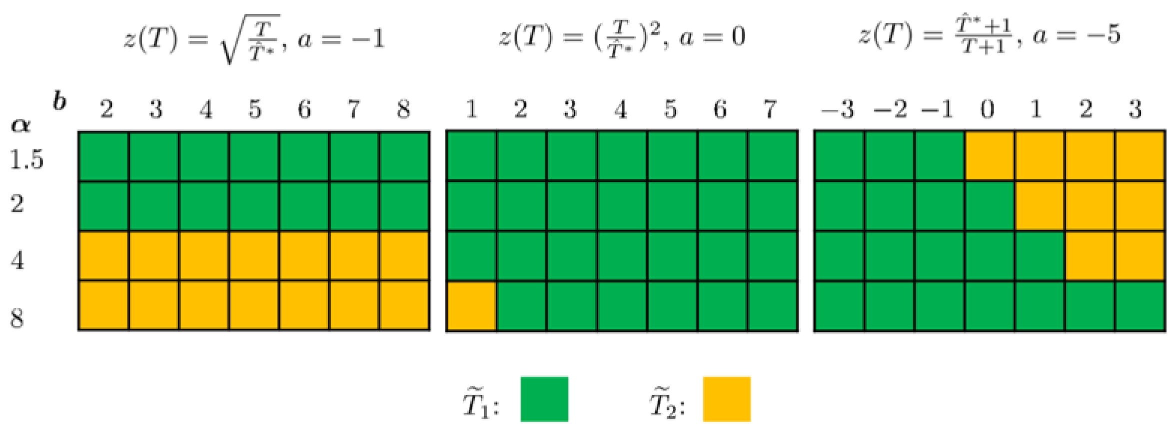


Figure 5: Preferred heuristic replacement age for different combinations of $z(T)$, $[a, b]$ and α where $X \sim \text{Weibull}(\alpha, 10)$ and $Y \sim \text{Uniform}[a, b]$. For each of the three cases of $z(T)$, the value of a is fixed and the preferred suboptimal solution is depicted for different values of α and b. The considered heuristics are $\widetilde{T}_1 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^* - \mu_Y)$ and $\widetilde{T}_2 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^*)$.

values of α. For $z(T) = \frac{\hat{T}^*+1}{T+1}$, prescribing replacement at $\widetilde{T}_2$ is less costly for larger values of b, i.e., larger mean and variance of Y. However, for larger values of α, $\widetilde{T}_1$ dominates.

Hence, we cannot draw any general conclusions about the superiority of one heuristic replacement age over the other. However, with insights generated by the analysis akin to that behind Figure 5, maintenance planners can make an informed choice of replacement age. For example, if the maintenance worker is always late and the variance of his/her degree of unpunctuality increases quickly over time, then the maintenance planner should opt for $\widetilde{T}_1$. However, if the variance of his/her degree of unpunctuality increases slowly over time, the planner should opt for $\widetilde{T}_1$ only if the failure rate of the system increases sufficiently slowly.

2.5.2 Cost-rate Bounds

First, we derive upper bounds on cost-rate ratios to characterize how non-stationary unpunctual policy implementation increases the cost-rate relative to punctual implementation (Proposition 6). Second, we compare the optimal cost-rates under non-stationary and stationary unpunctual implementation and provide a lower bound on the degree to which the maintenance planner can reduce the cost-rate by anticipating the non-stationary behavior (Proposition 7). Lastly, we examine how disregarding non-stationary unpunctual behavior, i.e., prescribing replacement in anticipation of punctual implementation, can increase the cost-rate (Proposition 8).

Proposition 6. *Assume $\widetilde{T}^*$ is the unique solution to problem* (2.5)*. If $\widetilde{T}_1 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^* - \mu_Y)$ and $\widetilde{T}_2 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^*)$ are feasible to problem* (2.7)*, then*

$$1 \leq \frac{\Omega(T^*)}{\widetilde{\Omega}(\widetilde{T}^*)} \leq \min\{U_Y^1(\widetilde{T}_1), U_Y^1(\widetilde{T}_2)\} \leq \min\{U^1(\widetilde{T}_1), U^1(\widetilde{T}_2)\} \tag{2.15}$$

where

$$\begin{aligned}
U_Y^1(T) =& \frac{\int_a^b \left(c_m \int_0^{T+yz(T)} h_X(x)dx + c_p\right) dF_Y(y)}{c_m h_X(T)(T + \mu_Y z(T))}, \textit{ and} \\
U^1(T) =& \frac{c_m M(T) + c_p}{c_m h_X(T)(T + \mu_Y z(T))}, \textit{ where} \\
& M(T) = \frac{\int_{T+az(T)}^{T+bz(T)} h_X(x)dx}{b-a}(\mu_Y - a) + \int_0^{T+az(T)} h_X(x)dx.
\end{aligned}$$

The first inequality in expression (2.15) implies that if the maintenance worker exhibits non-stationary unpunctual behavior, then the optimal cost-rate is higher compared to the case in which the maintenance worker is punctual (recall Theorem 2). Proposition 6 provides upper bounds on this increase in cost-rate. The maintenance planner can compute these bounds using the optimal replacement time under punctual implementation (i.e., $\widetilde{T}^*$). The presence of the subscript Y in (2.15) emphasizes the tighter bound's dependence on the full characterization of Y. In contrast, $U^1(T)$ only requires minimal knowledge of Y, i.e., a, b and μ_Y. If $U^1(T)$ is sufficiently tight, then the maintenance planner does not necessarily need to estimate the full distribution of the maintenance worker's unpunctuality.

If we assume that $z(\hat{T}^*) = 1$, then by Theorem 2

$$1 \leq \frac{\Omega(T^*)}{\widetilde{\Omega}(\widetilde{T}^*)} \leq \min\{\hat{U}_Y(\widetilde{T}^*), U_Y^1(\widetilde{T}_1), U_Y^1(\widetilde{T}_2)\} \leq \min\{U^1(\widetilde{T}_1), U^1(\widetilde{T}_2)\}$$

where

$$\hat{U}_Y(\widetilde{T}^*) = \frac{\int_a^b \left(c_m \int_0^{\widetilde{T}^*-\mu_Y+y} h_X(x)dx + c_p\right) dF_Y(y)}{c_m h_X(\widetilde{T}^*)\widetilde{T}^*}.$$

In [64], the cost-rate ratio $\frac{\hat{\Omega}(\hat{T}^*)}{\widetilde{\Omega}(\widetilde{T}^*)}$ is bounded by $\hat{U}_Y(\widetilde{T}^*)$, which is obtained using the sub-optimal solution $\widetilde{T}^* - \mu_Y$. In Proposition 6 we also use sub-optimal solutions, namely $\widetilde{T}_1$ and $\widetilde{T}_2$, to derive upper bounds. Depending on the model parameters and $z(T)$, either one of these sub-optimal replacement ages may yield a tighter bound than the other. Note that if we let $z(T) = 1$ for all T, then $U_Y^1(T)$ and $U^1(T)$ reduce to the bounds established in [64, see Theorem 4].

Table 2: Numerical example of the bounds in Propositions 6-8 for $X \sim$ Weibull(α, β) and $Y \sim$ Uniform$[a, b]$, $c_m = 1$ and $c_p = 16$, $z(T) = \exp(-\frac{T}{\hat{T}^*} + 1)$. Note that $U_Y^1 = \min\{U_Y^1(\widetilde{T}_1), U_Y^1(\widetilde{T}_2)\}$, $U^1 = \min\{U^1(\widetilde{T}_1), U^1(\widetilde{T}_2)\}$, $L_Y^2 = \min\{L_Y^2(\widetilde{T}_1), L_Y^2(\widetilde{T}_2)\}$ and $T_s^* = \text{argmin}\{\Omega(\widetilde{T}_1), \Omega(\widetilde{T}_2)\}$.

										I	II	III	IV	V	VI	VII	VIII	IX	X	XI
α	β	a	b	μ_Y	Var(Y)	$\widetilde{T}^*$	$\hat{T}^*$	T^*	T_s^*	$\frac{\Omega(T^*)}{\widetilde{\Omega}(\widetilde{T}^*)}$	$\hat{U}_Y(\widetilde{T}^*)$	U_Y^1	U^1	$\frac{\Omega(T^*)}{\hat{\Omega}(\hat{T}^*)}$	$L_Y^1(\widetilde{T}^*)$	$\frac{\Omega(\widetilde{T}^*)}{\Omega(T^*)}$	L_Y^2	$U_Y^2(\widetilde{T}^*)$	$U^2(\widetilde{T}^*)$	$\frac{\Omega(T_s^*)}{\Omega(T^*)}$
5	10	0	5	2.5	2.08	13.20	10.54	10.87	9.98	1.0231	1.0242	1.0233	1.0693	0.9993	0.9764	1.0389	1.0387	1.0629	1.1013	1.0002
6	10	0	5	2.5	2.08	12.14	9.39	9.80	8.95	1.0336	1.0362	1.0337	1.1026	0.9985	0.9650	1.0579	1.0578	1.0934	1.1551	1.0001
5	10	-4	0	-2	1.33	13.20	15.09	15.60	15.9	1.0152	1.0154	1.0152	1.0459	0.9999	0.9848	1.0577	1.0577	1.0738	1.1015	1.0000
6	10	-4	0	-2	1.33	12.14	13.98	14.05	14.86	1.0224	1.0230	1.0225	1.0682	0.9999	0.9775	1.0772	1.0771	1.1013	1.1349	1.0001
5	10	-4	5	0.5	6.75	13.20	12.20	12.81	12.52	1.0735	1.0802	1.0736	1.2252	0.9964	0.9257	1.0014	1.0013	1.0750	1.2240	1.0001
6	10	-4	5	0.5	6.75	12.14	10.87	11.57	11.48	1.1055	1.1240	1.1057	1.3312	0.9934	0.8897	1.0044	1.0043	1.1104	1.3369	1.0001

Table 3: Numerical examples of the bounds in Propositions 6-8 for $X \sim$ Weibull(α, β) and $Y \sim$ Uniform$[a, b]$, $c_m = 1$ and $c_p = 16$. Note that U_Y^1 =min$\{U_Y^1(\widetilde{T}_1), U_Y^1(\widetilde{T}_2)\}$, U^1 =min$\{U^1(\widetilde{T}_1), U^1(\widetilde{T}_2)\}$, L_Y^2 =min$\{L_Y^2(\widetilde{T}_1), L_Y^2(\widetilde{T}_2)\}$ and T_s^* =argmin$\{\Omega(\widetilde{T}_1), \Omega(\widetilde{T}_2)\}$.

							I	II	III	IV	V	VI	VII	VIII	IX	X	XI
	α	β	$\widetilde{T}^*$	$\hat{T}^*$	T^*	T_s^*	$\frac{\Omega(T^*)}{\widetilde{\Omega}(\widetilde{T}^*)}$	$\hat{U}_Y(\widetilde{T}^*)$	U_Y^1	U^1	$\frac{\Omega(T^*)}{\hat{\Omega}(\hat{T}^*)}$	$L_Y^1(\widetilde{T}^*)$	$\frac{\Omega(\widetilde{T}^*)}{\Omega(T^*)}$	L_Y^2	$U_Y^2(\widetilde{T}^*)$	$U^2(\widetilde{T}^*)$	$\frac{\Omega(T_s^*)}{\Omega(T^*)}$
$z(T) = (\frac{T}{\hat{T}^*})^{1.5}$, $a = 0$, $b = 5$	5	10	13.20	10.54	10.39	10.64	1.0234	1.0242	1.0242	1.0656	0.9995	0.9764	1.1933	1.1916	1.2212	1.3756	1.0014
	1.5	1	10.08	7.73	7.55	7.65	1.0049	1.0051	1.0050	1.0149	0.9998	0.9949	1.0256	1.0255	1.0306	1.0449	1.0001
$z(T) = (\frac{T}{\hat{T}^*})^{1.5}$, $a = -4$, $b = 5$	5	10	13.20	12.20	11.67	12.63	1.0723	1.0802	1.0802	1.2733	0.9953	0.9257	1.0393	1.0235	1.1145	1.3449	1.0154
	1.5	1	10.08	7.73	7.55	7.65	1.0141	1.0167	1.0145	1.0436	0.9980	0.9835	1.0020	1.0016	1.0162	1.0476	1.0003
$z(T) = 3\frac{T}{\hat{T}^*} - 2$, $a = -4$, $b = 0$	5	10	13.20	15.09	14.52	15.24	1.0136	1.0136	1.0136	1.0431	0.9983	0.9848	1.0091	1.0090	1.0228	1.0326	1.0000
	1.5	1	10.08	12.18	11.15	11.04	1.0024	1.0033	1.0024	1.0061	0.9992	0.9967	1.0011	1.0010	1.0035	1.0053	1.0000

Tables 2 and 3 illustrate the performance of the bounds established in Proposition 6 for monotone decreasing and monotone increasing $z(T)$, respectively. For all instances, we assume a form of $z(T)$ for which $z(\hat{T}^*) = 1$. Thus, we also include the upper bound $\hat{U}_Y(\widetilde{T}^*)$ established in [64] in column II of Tables 2 and 3. In both tables, the time to failure distribution is assumed to be Weibull with shape parameter α and scale parameter β. For these examples, most bounds are tight, hence they can be used to determine when unpunctual replacement increases the cost-rate significantly. For example, unpunctual replacements could cost the maintenance planner upwards of 10.55% (column I Table 2). With only μ_Y specified, the upper bound indicates an increase of 33.12% (column IV Table 2); with full information on Y, the corresponding upper bound is 10.57% (column III Table 2).

As seen in Tables 2 and 3, if the variance of the degree of unpunctuality is sufficiently large and the hazard rate of the system increases quickly (i.e., large α), then the increase in the cost-rate is larger and so are the bounds. Furthermore, U^1 is a looser bound than U_Y^1 in these instances compared to instances with smaller α and variance of Y. Also, note that in most instances, U_Y^1 outperforms $\hat{U}_Y$. The fact that U_Y^1 appears to be a tighter bound indicates that, compared to $\hat{T}^*$, the suboptimal solutions $\widetilde{T}_1$ and $\widetilde{T}_2$ (which incorporate the non-stationary behavior of the maintenance worker, i.e., they depend on $z(T)$) tend to be closer to T^*.

Next, Proposition 7 provides a lower bound on the ratio of the long-run cost-rates for problems (2.6) and (2.7) under condition (2.11). The bound in Proposition 7 gives a potential percent decrease in the cost-rate if the unpunctuality is non-stationary. Recall that Theorem 2 establishes that by anticipating the non-stationary behavior, replacement can be prescribed such that it reduces the mean and variance of the degree of unpunctuality, resulting in a smaller optimal cost-rate.

Proposition 7. *Assume $\hat{T}^*$ is feasible to problem* (2.7)*, and $\widetilde{T}^* - \mu_Y$ is feasible to problem* (2.6)*. If $z(\hat{T}^*) = 1$, then*

$$L_Y^1(\widetilde{T}^*) \le \frac{\Omega(T^*)}{\hat{\Omega}(\hat{T}^*)} \le 1, \tag{2.16}$$

where

$$L_Y^1(\widetilde{T}^*) = \frac{c_m h_X(\widetilde{T}^*)\widetilde{T}^*}{\int_a^b \left(c_m \int_0^{\widetilde{T}^*-\mu_Y+y} h_X(x)dx + c_p\right)dF_Y(y)}.$$

The second inequality in expression (2.16) is a direct result of Theorem 2. For the lower bound in (2.16), the characterization of $z(T)$ is not necessary. In the numerical results in Tables 2 and 3, the maximum percent decrease in optimal cost-rate ratio is 11.03% (column VI Table 2). For smaller values of α and smaller variance of Y, the lower bound on this decrease is tighter. Note that in the absence of the assumption $z(\hat{T}^*) = 1$, Proposition 7 still holds for cases in which $\Omega(T^*) \leq \hat{\Omega}(\hat{T}^*)$. More discussion on this comparison is provided in Section 2.6.

Lastly, we assess how prescribing replacement assuming punctual implementation can affect the cost-rate when, in fact, the implementation is unpunctual and depends on time.

Proposition 8. *Assume $\widetilde{T}^*$ is a unique solution to problem* (2.5) *and feasible to problem* (2.7)*. If $\widetilde{T}_1 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^* - \mu_Y)$ and $\widetilde{T}_2 = \widetilde{T}^* - \mu_Y z(\widetilde{T}^*)$ are feasible to problem* (2.7)*, then*

$$\max\{L_Y^2(\widetilde{T}_1), L_Y^2(\widetilde{T}_2), 1\} \leq \frac{\Omega(\widetilde{T}^*)}{\Omega(T^*)} \leq U_Y^2(\widetilde{T}^*) \leq U^2(\widetilde{T}^*), \tag{2.17}$$

where

$$\begin{aligned}
L_Y^2(T) =& \frac{\int_a^b \left(c_m \int_0^{\widetilde{T}^*+yz(\widetilde{T}^*)} h_X(x)dx + c_p\right)dF_Y(y)}{\int_a^b \left(c_m \int_0^{T+yz(T)} h_X(x)dx + c_p\right)dF_Y(y)} \cdot \frac{T + \mu_Y z(T)}{\widetilde{T}^* + \mu_Y z(\widetilde{T}^*)} \\
U_Y^2(\widetilde{T}^*) =& \frac{\int_a^b \left(c_m \int_0^{\widetilde{T}^*+yz(\widetilde{T}^*)} h_X(x)dx + c_p\right)dF_Y(y)}{c_m h_X(\widetilde{T}^*)(\widetilde{T}^* + \mu_Y z(\widetilde{T}^*))}, \text{ and} \\
U^2(\widetilde{T}^*) =& \frac{c_m M(\widetilde{T}^*) + c_p}{c_m h_X(\widetilde{T}^*)(\widetilde{T}^* + \mu_Y z(\widetilde{T}^*))}.
\end{aligned}$$

Proposition 8 gives lower and upper bounds on the ratio of the long-run cost-rates under non-stationary unpunctual implementation when replacement is scheduled at $\widetilde{T}^*$ and T^*, respectively. These bounds measure the loss associated with the maintenance planner ignoring the possibility of an unpunctual maintenance worker. Similar to Proposition 6, the subscript Y used in $U_Y^2(\widetilde{T}^*)$ shows the bound's dependence on the full characterization of Y, whereas to calculate $U^2(\widetilde{T}^*)$, only minimal knowledge of Y is required.

For the examples in Tables 2 and 3, if the maintenance planner disregards the non-stationary unpunctual behavior, the increase in cost-rate could be as large as 19.3% (column VII Table 3). With only μ_Y specified, the upper bound indicates an increase of 37.56% (column X Table 3); with full information on Y, the increase in the long-run cost-rate is as large as 22.12% (column IX Table 3). Moreover, similar to the bounds in Propositions 6 and 7, the upper bounds in Proposition 8 are rather loose for the instances with larger variance in the degree of unpunctuality.

2.6 A Numerical Example for $z(\hat{T}^*) \neq 1$ & $z(T^*) \neq 1$

In this section, we relax the assumption that $z(\hat{T}^*) = 1$ or $z(T^*) = 1$ and compare the optimal cost-rates of problems (2.5), (2.6) and (2.7) numerically. Example 7 does so assuming that $z(T) = T/d$, and varying value of d.

Example 7. Let $X \sim$ Weibull$(3, 1)$, $Y \sim$ Uniform$[0, 2]$ and $k = 4$. Then, $\widetilde{\Omega}^* = 4.76$, $\widetilde{T}^* = 1.26$, $\hat{\Omega}^* = 5.76$ and $\hat{T}^* = 0.26$. Table 4 presents the optimal cost-rates and replacement times under non-stationary unpunctuality with $z(T) = \frac{T}{d}$, where $d \in \{0.02, 0.1, 0.12, 0.15, 0.2, 0.26, 0.4, 10, 100\}$.

Recall from Theorem 2 that for $z(\hat{T}^*) = 1$, the optimal cost-rate under non-stationary unpunctual behavior is smaller than that of stationary unpunctual behavior. In Table 4, this relationship holds for values of $d \geq \hat{T}^*$ as well as $d = 0.2$ and $d = 0.15$. This result holds because under those scenarios, the mean and variance of unpunctuality at T^* is smaller under non-stationary unpunctual behavior compared to its stationary counterpart. Note that for $d = 0.12$ and $d = 0.1$, even though the same relationship holds (i.e., mean and variance of unpunctuality at T^* is smaller under non-stationary unpunctual behavior compared to its stationary counterpart), the optimal cost-rate is larger. In those scenarios, because the expected cycle length is very small under non-stationary unpunctual behavior, the expected cost-rate is larger compared to that under stationary counterpart. ■

As demonstrated by Table 4, in general it is hard to compare the optimal policies under non-stationary and stationary unpunctual behavior. Hence, making an assumption such as

Table 4: Optimal cost-rate and replacement time under non-stationary unpunctuality for Example 7 with $z(T) = \frac{T}{d}$, $d \in \{0.02, 0.1, 0.12, 0.15, 0.2, 0.26, 0.4, 10, 100\}$, and optimal cost-rate ratios $\Omega(T^*)/\hat{\Omega}(\hat{T}^*)$ and $\Omega(T^*)/\widetilde{\Omega}(\widetilde{T}^*)$. The column with $\hat{T}^* = 0.26$ corresponds to the case where $z(\hat{T}^*) = 1$; hence, the comparison of the optimal cost-rates and replacement times under stationary and non-stationary unpunctual behavior correspond to Theorems 2 and 5, respectively.

$\boldsymbol{d}$	0.02	0.1	0.12	0.15	0.2	$\boldsymbol{\hat{T}^* = 0.26}$	0.4	10	100
$\boldsymbol{T^*}$	0.02	0.09	0.11	0.14	0.17	0.22	0.3	1.14	1.25
$\boldsymbol{\Omega(T^*)}$	5.96	5.82	5.79	5.75	5.67	5.60	5.12	4.77	4.76
$\boldsymbol{\Omega(T^*)/\hat{\Omega}(\hat{T}^*)}$	1.03	1.01	1.005	0.99	0.98	0.97	0.89	0.83	0.82
$\boldsymbol{\Omega(T^*)/\widetilde{\Omega}(\widetilde{T}^*)}$	1.25	1.22	1.226	1.21	1.19	1.18	1.08	≈ 1	≈ 1

$z(\hat{T}^*) = 1$ or $z(T^*) = 1$ is helpful for deriving analytical results. Nevertheless, in most scenarios, when comparing two maintenance workers with non-stationary and stationary behavior, if $z(\hat{T}^*)$ or $z(T^*)$ are sufficiently close to 1, then the results of Theorems 2 and 3 apply, respectively. Moreover, in most cases, a maintenance worker with sufficiently larger degree of unpunctuality increases the cost-rate compared to a maintenance worker with smaller degree of unpunctuality.

2.7 Concluding Remarks

We consider a novel age-replacement policy with minimal repair that optimally anticipates potential deviations between prescribed and actual replacement times when the distribution of this deviation depends on the prescribed replacement time itself. We consider both the possibility that the maintenance worker becomes more punctual as replacement is scheduled further into the future, as well as the possibility that the maintenance worker becomes less punctual as replacement is scheduled further into the future.

We compare the optimal planned replacement time for this problem to that for the case in which the nature of the unpunctual behavior of the worker is constant over time, i.e., independent of the prescribed replacement time. A comparison of the optimal replacement ages reveals that if the mean and variance of the degree of deviation from the prescribed replacement time are increasing (decreasing), then it is optimal to prescribe replacement earlier (later) than we would for the stationary unpunctual case. This result suggests that the optimal policy aims to decrease the expected degree of unpunctuality and its variance by adjusting the optimal planned replacement time.

Furthermore, we compare these optimal replacement times and their cost-rates to those for the case in which the maintenance worker is punctual. Under both stationary and non-stationary unpunctuality, the optimal cost-rate is greater than it is for the punctual case. However, in our numerical examples the optimal cost-rate for our problem is less than that under stationary unpunctual behavior when the deviation distributions are similar near the optimal replacement time for the stationary case; we prove that this result holds without exception if distributions coincide at the optimal replacement time for the stationary case. This decreased cost-rate reflects the benefit of capitalizing on the non-stationary behavior to reduce the mean and variance of the degree of unpunctuality by adjusting the planned replacement time.

We also provide a lower bound to characterize how much the maintenance planner can gain under non-stationary behavior relative to the stationary case, as well as upper bounds to characterize how much they can lose by ignoring non-stationary unpunctuality and assuming punctual behavior. These bounds do not require the optimal solution to our problem. Moreover, we provide upper bounds that do not require the full knowledge of the distribution of the unpunctual behavior. Therefore, these bounds can reflect the value of estimating the true distribution of the unpunctual behavior.

Finally, we present scenarios in which the mean and variance of the degree of unpunctuality are smaller under non-stationary unpunctual behavior compared to stationary behavior, but the expected cost-rate is larger. Hence, for some instances, even though the optimal policy is able to reduce the mean and variance of unpunctuality by replacing early or late, it cannot reduce the overall cost-rate compared to the scenario when the worker behaves consistently. Therefore, for such instances it may be valuable to incentivize workers to be more consistent or punctual.

3.0 Optimal Age-Replacement under Time-Dependent Replacement Costs

3.1 Introduction

Age-replacement with minimal repair after failure under age-dependent replacement costs has been well studied in the literature. Authors in [24] and [25] consider increasing minimal repair costs and prove existence and uniqueness of the optimal policy. Another study in [127] defines an increasing cost function $a(t)$, where t denotes system age, add $a(t)$ to the long-run expected total cost function, and prove existence and uniqueness of the optimal policy under various functional forms of $a(t)$. Authors in [118] assume non-decreasing minimal repair and preventive replacement costs, formulate a semi-Markov decision process model, and establish that the optimal policy is of threshold type under mild assumptions. Lastly, the study in [120] consider age-dependent minimal repair costs for a k-out-of-n system paired with ordering decisions.

In this chapter, we focus on age-replacement *without minimal repair* under age-dependent replacement costs, which is less well studied in literature, and to the best of our knowledge, is limited to the studies in [9, 39, 116]. The model developed in [116] incorporates a cost factor that increases with the age of the system and is proportional to the length of time between two replacements (when the system is functioning). This study provides conditions for the existence and uniqueness of the optimal policy for specific forms of the cost factor (polynomial and exponential functions of the expected cycle length) and exponential time to failure distribution. Our existence and uniqueness results are more general in that we do not assume specific forms of the cost factor and the time to failure distribution.

Moreover, the authors in [9] consider age-replacement without minimal repair under age-dependent replacement costs, but in conjunction with spare part ordering decisions. They assume that at most one spare part can be kept in stock or on order at each time epoch, and restrict their attention to the case in which the reactive and preventive replacement costs differ by a constant and are convex in age. We consider a variant of this model, namely that in which a spare unit is always available. However, we allow the replacement cost functions to exhibit different growth rates and take either convex or concave forms. Concave forms are motivated by the observation that replacement costs can increase at a decreasing rate as equipment ages (see e.g., [77, 81]), and require novel conditions to ensure the unimodality of the cost-rate function (Section 3.3).

Authors in [39] considers age-replacement without minimal repair for products that are sold with a warranty policy. In their model, if the product fails during the warranty period, then a refund is received in an amount proportional to the failure time; if it fails after the warranty period, then no refund is received and a constant reactive replacement cost is assumed. If the product is preventively replaced, then it is salvaged at a value that is linearly proportional to its residual lifetime. That is, the reactive replacement cost after the warranty period is not age-dependent and the functional form of the preventive replacement cost is restricted to a linear form.

Our contributions are summarized as follows. We generalize the literature on age-replacement without minimal repair by allowing for age-dependent replacement cost functions to take on more general functional forms (Section 3.2). We provide conditions under which there exists a unique optimal solution under a long-run expected cost-rate minimization objective (Section 3.3). We compare the optimal replacement and long-run expected cost-rates under age-dependent and constant replacement costs (Section 3.4). We then generalize to non-instantaneous replacements and consider an availability criterion (Section 3.5). Finally, we provide a summary of our findings and discuss future research directions (Section 3.6). The proofs for all results are provided in Appendix B.

3.2 Model Formulation

Consider a stochastically deteriorating system with self-announcing failures that require immediate reactive replacement. Here, we assume that replacements are instantaneous. Let the continuous random variable X be the time to failure of the system, with c.d.f. $F(x)$ for $x \geq 0$, p.d.f. $f(x)$, survival function $\bar{F}(x)$, hazard rate function $h(x)$, i.e., $h(x) = \frac{f(x)}{\bar{F}(x)}$, and mean μ ($\lim_{t\to\infty} \int_0^t \bar{F}(x)dx = \mu$). We assume that $F(0) = 0$ and impose the following conditions on $h(x)$:

Assumption 4. $h(0) = 0$ *and* $h(x)$ *is strictly increasing to* $+\infty$*.*

Note that the notations we use in this chapter are somewhat consistent with those in Chapter 2, and Assumption 4 impose the same conditions as in Assumption 1. However, the two models are different in that, here, the system is replaced upon failure (i.e., preventive maintenance without minimal repair) whereas in Chapter 2 the system is minimally repaired.

For such deteriorating systems, age-replacement policies are commonly used in practice to offset the replacement costs incurred by system failures [131]. Here, we consider an age-replacement policy without minimal repair, and denote this policy by T. That is, the system is replaced T units of time after its installation (i.e., when it ages T units of time) or at failure, whichever occurs first. We have $0 < T \leq +\infty$ where $T = +\infty$ indicates that preventive replacement is never optimal. Motivated by the applications described in the previous section, we assume that upon replacement the system is exchanged with one that is as-good-as-new and has the same age-dependent replacement cost functions.

We denote the age-dependent reactive replacement cost at time t as $c_r(t)$ and the preventive replacement cost at scheduled time t as $c_p(t)$. Throughout this chapter, we impose the following assumption on the cost functions $c_r(t)$ and $c_p(t)$:

Assumption 5. *For all* t*,* $c_r(t) - c_p(t) > \delta$ *for some* $\delta > 0$*. Also,* $c_r(t)$ *and* $c_p(t)$ *are continuous, three times differentiable, and* $c_r'(t) \geq 0$ *and* $c_p'(t) \geq 0$ *for all* t*.*

The first condition in Assumption 5 implies that the cost of reactive replacement is strictly larger than the cost of preventive replacement at all times. This assumption is common in the maintenance literature [13]. The assumptions of continuity and differentiability are made for mathematical convenience. Note that constant replacement costs satisfy these assumptions. Finally, we assume replacement costs are non-decreasing in system age.

Let the sequence of random variables $\{R_1, R_2, ...\}$ denote the replacement times of the system, and the sequence of random variables $\{\mathcal{X}_1, \mathcal{X}_2, ...\}$ denote the times to failure, where $\mathcal{X}_i$ has p.d.f. $f(x)$. Then, $R_i - R_{i-1} = \min\{T, \mathcal{X}_i\}$ for $i \in \{1, 2, ...\}$ where $R_0 = 0$. Let ζ_i denote the replacement cost incurred at the ith replacement time. Then, under age-replacement policy T,

$$\zeta_i = \begin{cases} c_p(R_i - R_{i-1}) & \text{if } \mathcal{X}_i \geq T \\ c_r(R_i - R_{i-1}) & \text{if } \mathcal{X}_i < T \end{cases} = \begin{cases} c_p(T) & \text{if } \mathcal{X}_i \geq T \\ c_r(\mathcal{X}_i) & \text{if } \mathcal{X}_i < T. \end{cases} \tag{3.1}$$

Equation (3.1) holds because the replacement costs depend on the system's age (i.e., the time elapsed since the last renewal). Figure 6 depicts the relationship between ζ_i and the replacement cost functions $c_p(t)$ and $c_r(t)$ under various forms of $c_p(t)$ and $c_r(t)$. Figure 6 also compares our setting with the conventional age-replacement model under constant replacement costs studied in [13].

Because replacement actions return the system to as-good-as-new, we can take a renewal-reward approach and obtain the long-run expected cost-rate as the ratio of the expected renewal cycle cost to the expected renewal cycle length. For an age-replacement policy T, let $\mathbb{E}_X[C(T)]$ and $\mathbb{E}_X[L(T)]$ be the expected cycle cost and length, respectively. Then, in this chapter we have

$$\begin{aligned} \mathbb{E}_X[C(T)] &= \int_0^T c_r(x) f(x) dx + c_p(T)\bar{F}(T), \text{ and} \\ \mathbb{E}_X[L(T)] &= \int_0^T \bar{F}(x) dx. \end{aligned}$$

Hence, the optimization problem for minimizing the long-run expected cost-rate under age-dependent replacement costs is given by

$$\min_{T>0} \Omega(T) = \frac{\mathbb{E}_X[C(T)]}{\mathbb{E}_X[L(T)]} = \frac{\int_0^T c_r(x) f(x) dx + c_p(T)\bar{F}(T)}{\int_0^T \bar{F}(x) dx}. \tag{3.2}$$

We denote the optimal solution to problem (3.2) by T^*. Throughout this chapter, we refer to $\Omega(T)$ as the "cost-rate" under policy T for brevity.

We study problem (3.2) under the assumption that the replacement costs are non-decreasing in age. We provide analytical results that guarantee the existence and uniqueness of the solution to problem (3.2). Moreover, Section 3.4 compares the optimal replacement policy and cost-rate under age-dependent replacement costs with those under constant replacement costs. Finally, Section 3.5 introduces non-instantaneous repair times and compares the optimal cost-rate minimizing and availability-maximizing replacement policies. Appendix A provides additional notation that is used in the proofs and reviews some previous results. Appendix B provides the proofs for all results established in Sections 3.3-3.5.

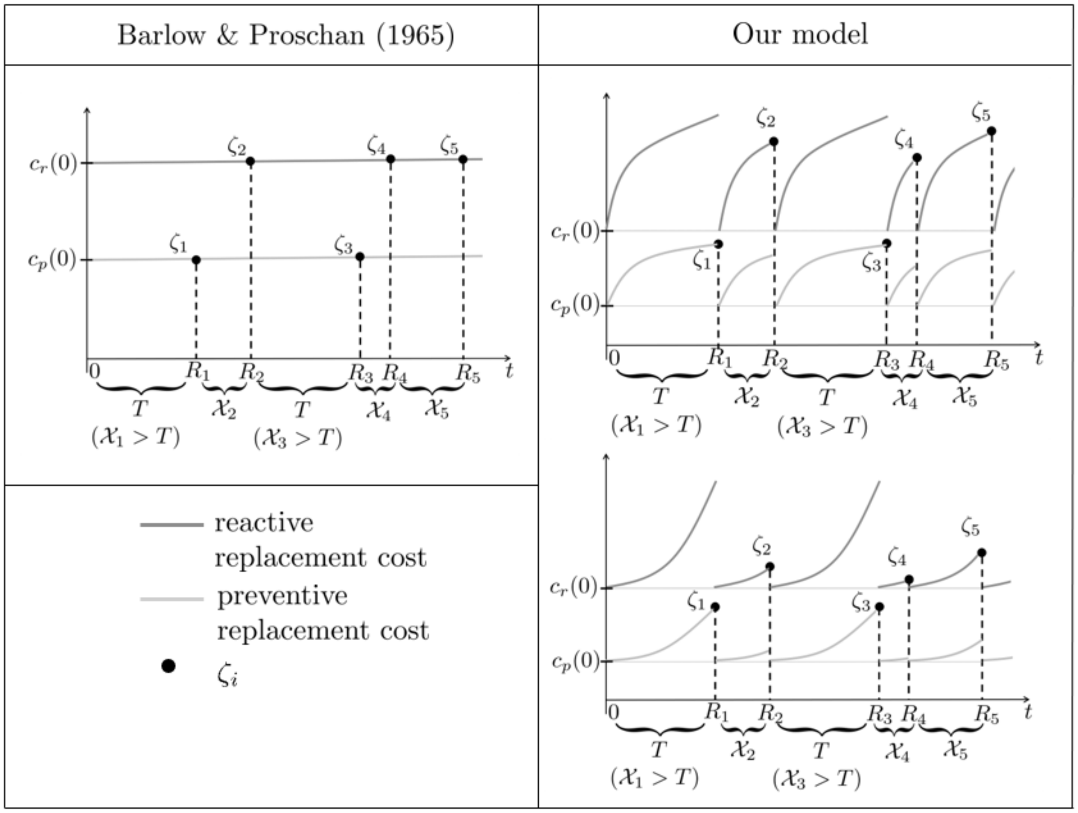


Figure 6: In each plot, a sequence of replacement times R_i is depicted for a given age-replacement policy T. The random variable ζ_i denotes the corresponding replacement cost incurred at time R_i. The plot on the left-hand side represents the case of constant replacement costs previously studied by [13]. The plots on the right-hand side represent the case of concave and convex non-decreasing replacement costs studied in this chapter.

3.3 Existence of a Unique Optimal Solution

Theorem 6 establishes sufficient conditions to ensure that there exists a unique finite optimal solution for problem (3.2).

Theorem 6. *Consider the following conditions: (i) $c_p''(t) \geq 0$, (ii) $c_r'(t) \geq c_p'(t)$, (iii) $c_r''(t) \geq c_p''(t)$, (iv) $c_p'''(t) \geq 0$ and (v) $h''(t) > 0$. If either conditions* (i) *and* (ii) *hold for all t, or* *conditions $(ii)-(v)$ hold for all t, then $\Omega(t)$ is quasi-convex and there exists a unique finite solution T^* to problem* (3.2)*. Moreover,* $\Omega(T^*) = \big(c_r(T^*) - c_p(T^*)\big)h(T^*) + c_p'(T^*)$.

Condition (i) requires $c_p(t)$ to be convex. Condition (ii) implies that the rate of increase in the reactive replacement cost is at least as large as that for the preventive replacement cost. That is, for example, the salvage value of a failed system decreases more quickly than that of a working system. Or, the efforts associated with replacing a failed system increase more quickly than a working system. Under these two conditions, the cost-rate function decreases to its unique minimum value and then increases.

Conditions $(ii)-(v)$ in Theorem 6 relax the convexity assumption of $c_p(t)$. Instead, they require the third derivative of $c_p(t)$ to be non-negative and $h(t)$ to be convex. The convexity of $h(t)$ eliminates scenarios under which the cost-rate is decreasing and preventive replacement is never optimal (i.e., $T^* = +\infty$). A positive third derivative eliminates scenarios under which the replacement cost function increases quickly (i.e., is convex) for a period of time, and then increases slowly (i.e., is concave). Example 8 considers an instance with non-convex $c_p(t)$ (i.e., condition (i) does not hold) in which condition (iv) is also violated (see Figure 7(a)) and as a result, the cost-rate is not quasi-convex (see Figure 7(b)).

Example 8. Let X follow a Weibull distribution with shape and scale parameter values of 3 and 50, respectively. Additionally, let $c_r(t) = \big(1 + \exp(-0.5(t-30))\big)^{-1} + 6$ and $c_p(t) = \big(1 + \exp(-0.5(t-30))\big)^{-1} + 3$. The corresponding replacement cost functions and long-run expected cost-rates are depicted in Figure 7.

In Example 8, the first local minimum, which is also the optimal replacement time, occurs at age $T^* = 27.41$; at this time, the cost of replacement is relatively smaller than it is at ages greater than 30. However, the cost-rate starts decreasing again at age 33.52, and reaches a second local minimum at age 44.01. This decrease in cost-rate occurs because under policy T_1, the expected cycle length is relatively longer than it is at ages less than 30, which balances the sudden increase in replacement costs. ■

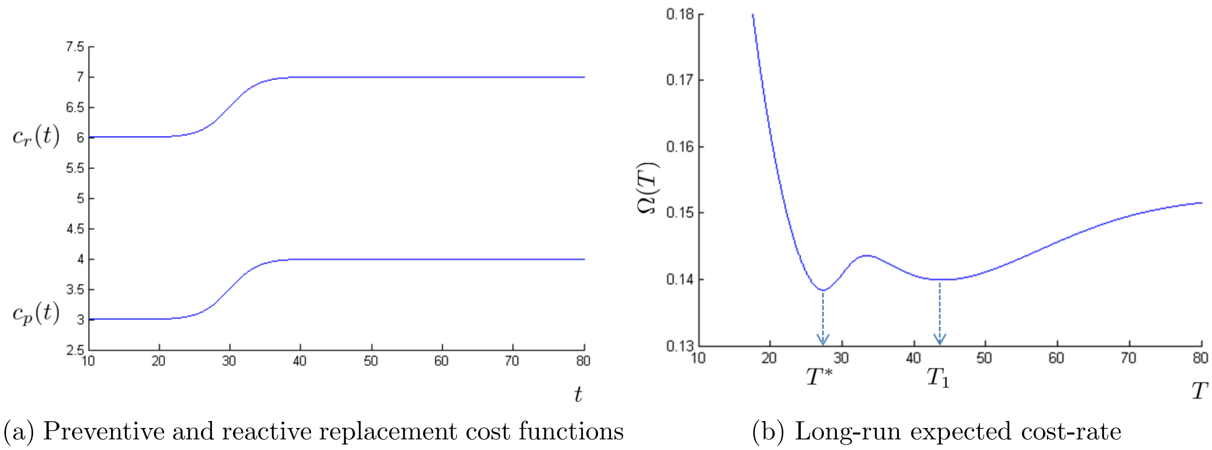


Figure 7: Under the parameter values of Example 8, the replacement cost functions violate condition (iv) in Theorem 6. As a result, the cost-rate function has two local minima at ages T^* and T_1.

In the remainder of this chapter, to ensure the existence and uniqueness of T^*, we make the following assumption.

Assumption 6. *Either conditions* $(i) - (ii)$ *or conditions* $(ii) - (v)$ *in Theorem 6 hold for all* t.

Assumption 6 is essential to derive analytical results that compare the optimal cost-rate minimizing or availability-maximizing replacement policies for different replacement costs.

3.4 Increasing vs. Constant Replacement Costs

In practice, replacement cost functions can be obtained by applying curve fitting techniques to historical data [23, 8]. In some settings, however, maintenance planners may choose to ignore the age-dependent nature of replacement costs when deriving replacement policies for their assets, either because assuming constant replacement costs may be attractive in terms of model simplification, or due to data sparsity. Prescribing preventive replacement policies according to constant replacement costs when in fact these costs increase in age, however, can increase the cost-rate significantly. This increase may be significant when replacement costs vary considerably in age, or when a company preventively maintains a large number of assets (recall our discussion in Section 3.1). Hence, analyzing the effect of age-dependent replacement costs on the optimal policy and the cost-rate can be insightful for maintenance planners.

For notational convenience, we denote the constant reactive and preventive replacement costs as $\widetilde{c}_r$ and $\widetilde{c}_p$, respectively. Moreover, we denote the long-run expected cost-rate under constant costs as $\widetilde{\Omega}(T)$. The corresponding optimization problem is given by

$$\min_{T>0} \widetilde{\Omega}(T) = \frac{\widetilde{c}_r F(T) + \widetilde{c}_p \bar{F}(T)}{\int_0^T \bar{F}(t)dt}, \tag{3.3}$$

the classical age-replacement problem without minimal repair [13]. We denote the optimal solution to problem (3.3) by $\widetilde{T}^*$. See Appendix A for a discussion on the existence of a unique finite solution to problem (3.3).

Section 3.4.1 provides analytical results and numerical examples that compare the optimal long-run expected cost-rate and the optimal replacement time under constant and age-dependent replacement costs. Section 3.4.2 examines the increase in the long-run expected cost-rate induced by ignoring age-dependent costs (i.e., $(\Omega(\widetilde{T}^*) - \Omega(T^*))/\Omega(T^*)$).

3.4.1 $\widetilde{\Omega}(\widetilde{T}^*)$ vs. $\Omega(T^*)$ and $\widetilde{T}^*$ vs. T^*

Theorem 7 compares the optimal policies under age-dependent and constant replacement costs under the assumption that they coincide at time zero and the growth rate of reactive and preventive replacement costs are identical.

Theorem 7. *Let $c_r(t) = \widetilde{c}_r + g(t)$ and $c_p(t) = \widetilde{c}_p + g(t)$, where $g(0) = 0$. (i) If $g(t)$ is convex, then $T^* \leq \widetilde{T}^*$. (ii) If $g(t)$ is concave, then $T^* \geq \widetilde{T}^*$. (iii) If $g(t) = at$ where $a \geq 0$, then $T^* = \widetilde{T}^*$ and $\Omega(t) - \widetilde{\Omega}(t) = a$ for all t. Under all these conditions, $\Omega(T^*) \geq \widetilde{\Omega}(\widetilde{T}^*)$.*

Under the conditions of Theorem 7, because reactive and preventive replacement costs have identical growth rates, their ratio decreases in system age. Nonetheless, when age-dependent replacement costs are convex, the optimal policy under age-dependent costs replaces more frequently than the optimal policy under constant costs. The opposite holds when replacement costs are concave. It is intuitive that $T^* \leq \widetilde{T}^*$ under convex replacement costs because replacement costs increase quickly as system ages.

Somewhat surprisingly, when age-dependent replacement costs have identical linear growth rates, T^* and $\widetilde{T}^*$ coincide even though $c_r(T^*)/c_p(T^*) < \widetilde{c}_r/\widetilde{c}_p$. To explain this result, recall from Section 3.2 the sequence of replacement times R_i, and their corresponding replacement costs ζ_i. Under the conditions of Theorem 7, for each i, for a given policy the difference between ζ_i under age-dependent and constant replacement costs is $a \cdot R_i$. Hence, the difference between the long-run expected cycle costs under age-dependent and constant replacement costs is $a \cdot \mathbb{E}_X[L(T)]$. Consequently, the long-run expected cost-rates differ by the constant a, and the optimal policies coincide.

Under the conditions of Theorem 7, the reactive and preventive replacement costs have identical growth rates. Proposition 9 and Theorem 8 relax this assumption and compare the replacement policies under different sets of conditions. Lastly in this section, Example 9 provides numerical examples that consider replacement costs with non-identical growth rates.

Proposition 9. *If $c_r(0) \geq \widetilde{c}_r$ and $c_p(\widetilde{T}^*) \leq \widetilde{c}_p$, then $T^* \leq \widetilde{T}^*$.*

Proposition 9 states that if reactive replacement is always more costly under the age-dependent case compared to its constant counterpart, but the preventive replacement at ages smaller than $\widetilde{T}^*$ is less costly than its constant counterpart, then it is optimal to replace more frequently.

Next, Theorem 8 compares the optimal policies under the special case where the replacement costs coincide at the optimal replacement time under constant replacement costs (i.e., $\widetilde{T}^*$).

Theorem 8. *If $c_r(\widetilde{T}^*) = \widetilde{c}_r$ and $c_p(\widetilde{T}^*) = \widetilde{c}_p$, then $T^* \leq \widetilde{T}^*$ and $\Omega(T^*) \leq \widetilde{\Omega}(\widetilde{T}^*)$.*

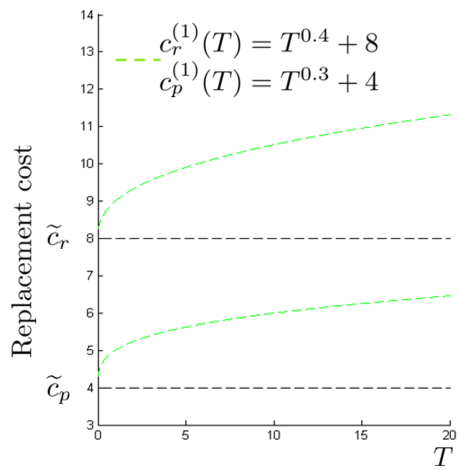
First, note that unlike Proposition 9, under the conditions of Theorem 8, the reactive replacement cost is smaller under the age-dependent case compared to its constant counterpart for $T^* \leq \widetilde{T}^*$ (i.e., $c_r(t) \leq \widetilde{c}_r$ for all $t \leq \widetilde{T}^*$ by Assumption 5). Intuitively, the result of Theorem 8 holds because by scheduling replacement earlier than $\widetilde{T}^*$, preventive replacement is less expensive under the age-dependent case compared to the constant case, and if the system fails before this age, then reactive replacement is also less expensive.

In general, without such conditions like those in Theorems 7, 8 and Proposition 9, it is not straightforward to make analytical comparisons. Example 9 examines the relationship between the optimal policies and cost-rates numerically for various replacement cost functions in the absence of such conditions.

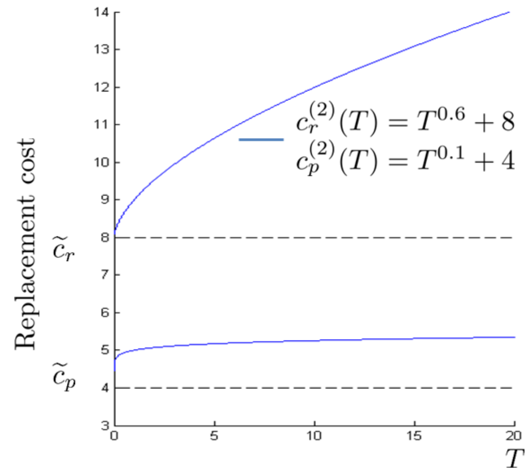
Example 9. Let X follow a Weibull distribution with shape and scale parameter values of 3 and 10, respectively. Four sets of replacement costs $\{c_r^{(1)}(T), c_p^{(1)}(T); c_r^{(2)}(T), c_p^{(2)}(T); c_r^{(3)}(T), c_p^{(3)}(T); c_r^{(4)}(T), c_p^{(4)}(T)\}$ are presented in Figure 8 and their corresponding cost-rates $\{\Omega^{(1)}(T); \Omega^{(2)}(T); \Omega^{(3)}(T); \Omega^{(4)}(T)\}$ are presented in Figure 9. Lastly, let $\widetilde{c}_r = 8$ and $\widetilde{c}_p = 4$ and the corresponding cost-rate function $\widetilde{\Omega}(T)$ is presented in Figure 9.

In Figures 8(a) and 8(b), age-dependent replacement costs are concave and coincide with their constant counterparts at age 0. In Figure 8(b), the ratio of age-dependent reactive and preventive replacement costs are larger than those in Figure 8(a) for larger values of system age. Hence, the optimal replacement policy under $\{c_r^{(2)}(T), c_p^{(2)}(T)\}$ is earlier than $\widetilde{T}^*$, whereas the optimal replacement policy under $\{c_r^{(1)}(T), c_p^{(1)}(T)\}$ is later than $\widetilde{T}^*$ (see Figure 9).

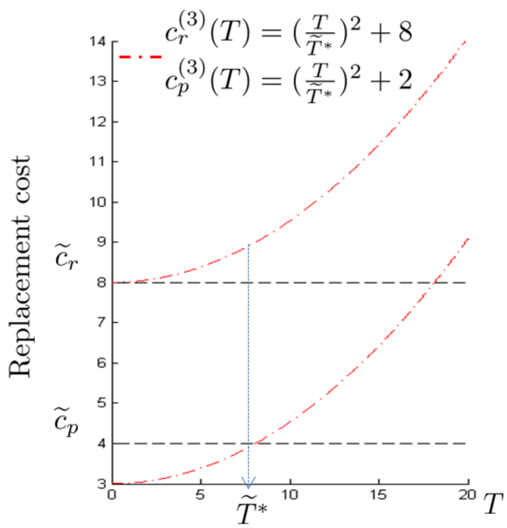
Moreover, the replacement costs in Figure 8(c) meet the conditions in Proposition 9, and hence the optimal replacement time under $\{c_r^{(3)}(T), c_p^{(3)}(T)\}$ is earlier than $\widetilde{T}^*$. For this scenario, the optimal cost-rate is smaller under the age-dependent costs because by replacing earlier than $\widetilde{T}^*$, the cost of preventive replacement is less than that under constant costs which offsets the higher cost of replacement at failure. Finally, under the convex replacement costs $\{c_r^{(4)}(T), c_p^{(4)}(T)\}$, the optimal replacement policy is earlier than $\widetilde{T}^*$ because the costs increase quickly, and their ratio also increases in system age. ■



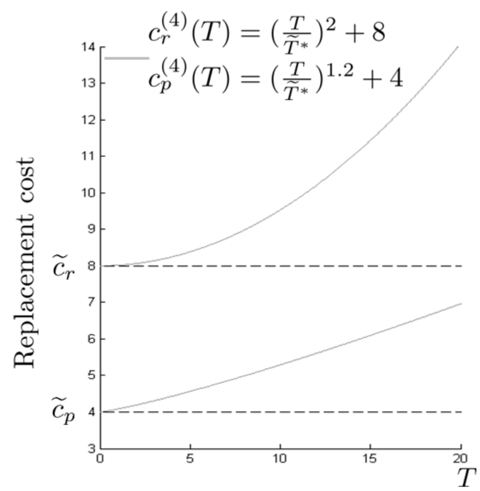
(a) Coincide at 0 and concave replacement costs



(b) Coincide at 0 and concave replacement costs



(c) $c_r^{(3)}(T) \geq \widetilde{c}_r \ \forall T$ & $c_p^{(3)}(\widetilde{T}^*) \leq \widetilde{c}_p$ (Proposition 9)



(d) Coincide at 0 and convex replacement costs

Figure 8: Reactive and preventive replacement cost functions in Example 9. Their corresponding long-run expected cost-rate functions are depicted in Figure 9.

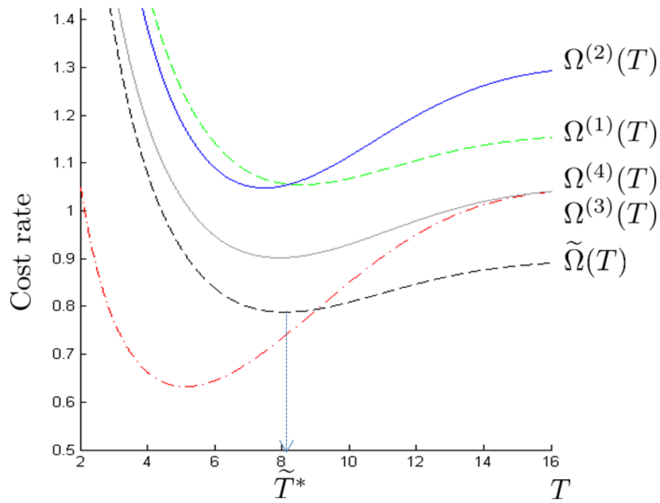


Figure 9: Cost-rate functions corresponding to the replacement cost functions in Figure 8 and Example 9. The replacement costs in $\Omega^{(3)}(T)$ meet the conditions in Proposition 9.

3.4.2 Price of Ignoring Age-dependency

In this section, we examine how much maintenance planners could lose if they prescribe replacement policies based on constant replacement costs, when in fact replacement costs depend on system age. We provide an example that considers various functional forms for the replacement cost functions, and compares the increase in long-run expected cost-rate induced by adopting replacement policies based on constant replacement costs.

Note that when reactive and preventive replacement costs are linear, with identical growth rates, prescribing a replacement policy based on constant replacement costs of a system at time zero does not affect the optimal policy; hence, the price of ignoring age-dependency is zero (recall Theorem 7). However, under more general forms of the replacement cost functions, the increase in the cost-rate can be large if the replacement policy is based on constant costs.

Example 10 illustrates the increase in the long-run expected cost-rate induced by adopting replacement policies $\widetilde{T}_0^*$, $\widetilde{T}_{\mu/2}^*$ and $\widetilde{T}_\mu^*$, where policy $\widetilde{T}_0^*$ is obtained by minimizing $\widetilde{\Omega}(T)$ under $\widetilde{c}_r = c_r(0)$ and $\widetilde{c}_p = c_p(0)$; $\widetilde{T}_{\mu/2}^*$ and $\widetilde{T}_\mu^*$ are obtained by minimizing $\widetilde{\Omega}(T)$ under

$\widetilde{c}_r = c_r(\mu/2)$ and $\widetilde{c}_p = c_p(\mu/2)$, and $\widetilde{c}_r = c_r(\mu)$ and $\widetilde{c}_p = c_p(\mu)$, respectively (recall that μ is the expected time to failure of the system).

Example 10. Let X follow a Weibull distribution with shape and scale parameter values 2.5 and 50, respectively, and let $\widetilde{c}_r = 3$ and $\widetilde{c}_p = 1$. Table 5 illustrates the percent increase in long-run expected cost-rate under $\widetilde{T}^*_0$, $\widetilde{T}^*_{\mu/2}$ and $\widetilde{T}^*_\mu$ for five scenarios. In all scenarios, $\widetilde{T}^*_0 = 32.84$ (because $c_r(0) = \widetilde{c}_r$ and $c_r(0) = \widetilde{c}_p$ under all scenarios).

In Example 10, the replacement costs in scenarios (a) and (b) are concave. Moreover, the ratio of the reactive and preventive replacement cost is decreasing in age. Hence, it is optimal to prescribe replacement policies later than $\widetilde{T}^*_0$; under scenarios (a) and (b), T^* is equal to 50.67 and 71.94, respectively. As a result, there is a substantial increase in the cost-rate if replacement policies are based on the replacement cost of a new system (i.e., $\widetilde{T}^*_0$). However, the price of ignoring age-dependency is less if replacement policies are based on the replacement cost of the system at $\mu/2$ or μ, because the ratio of the reactive and preventive replacement costs is smaller at ages $\mu/2$ and μ compared to 0.

Under scenarios (c) and (d), again, the ratio of the reactive and preventive replacement cost is decreasing in age. However, because the replacement costs increase more quickly compared to the concave scenarios, it is optimal to prescribe replacement policies earlier than $\widetilde{T}^*_0$; under scenarios (c) and (d), T^* is equal to 22.03 and 21.65, respectively. As a result, under $\widetilde{T}^*_0$, the increase in cost-rate is smaller than that under $\widetilde{T}^*_{\mu/2}$ and $\widetilde{T}^*_\mu$. Under scenario (e), the optimal replacement time is 17.55, and the ratio of the reactive and preventive replacement cost is first decreasing in age and then increasing. Hence, the increase in cost-rate is small for large ratios of reactive and preventive replacement costs. For example, the increase in cost-rate under the policy based on $\widetilde{c}_r = c_r(3\mu)$ and $\widetilde{c}_p = c_p(3\mu)$ is 0.36%. ■

From Example 10, we can conclude that assuming constant replacement costs, when in fact they are age-dependent, can increase the cost-rate significantly, even when the rate of increase in replacement costs is not large. Nevertheless, maintenance planners may favor heuristic policies based on constant replacement costs. The results of Example 10 suggest several insights. First, examining the rate of change in the ratio of the reactive and replacement costs is important in choosing a heuristic policy based on constant replacement costs. Moreover, if replacement costs are concave, then opting for a replacement policy based

Table 5: Percent increase in long-run expected cost-rate induced by $\widetilde{T}_0^*$, $\widetilde{T}_{\mu/2}^*$ and $\widetilde{T}_\mu^*$ under different age-dependent replacement costs in Example 10.

Scenario	Replacement costs	$\frac{\Omega(\widetilde{T}_0^*)-\Omega(T^*)}{\Omega(T^*)}$	$\frac{\Omega(\widetilde{T}_{\mu/2}^*)-\Omega(T^*)}{\Omega(T^*)}$	$\frac{\Omega(\widetilde{T}_\mu^*)-\Omega(T^*)}{\Omega(T^*)}$
(a)	$c_r(t) = t^{0.2} + \widetilde{c}_r$ $c_p(t) = t^{0.2} + \widetilde{c}_p$	7.48%	0.02%	0.12%
(b)	$c_r(t) = \log(5t+1) + \widetilde{c}_r$ $c_p(t) = \log(5t+1) + \widetilde{c}_p$	17.98%	0.0001%	0.03%
(c)	$c_r(t) = t/4 + \widetilde{c}_r$ $c_p(t) = t/8 + \widetilde{c}_p$	6.08%	14.21%	16.35%
(d)	$c_r(t) = (1.04)^t + \widetilde{c}_r - 1$ $c_p(t) = (1.04)^t + \widetilde{c}_p - 1$	8.47%	31.85%	66.80%
(e)	$c_r(t) = (1.07)^t + \widetilde{c}_r - 1$ $c_p(t) = (1.05)^t + \widetilde{c}_p - 1$	29.45%	65.46%	44.84%

on the costs at time $\mu/2$ may yield a small increase in cost-rate. However, if replacement costs are convex, then opting for a replacement policy based on the costs at time 0 may yield a small increase in cost-rate. Based on our numerical study, these conclusions hold across a wide range of values of the shape and scale parameters for the Weibull distribution.

3.5 Non-Instantaneous Replacements

In Sections 3.2-3.4, we assume that replacements are instantaneous; hence, the system is effectively available at all times. In this section, we generalize optimization problem (3.2) by allowing for non-instantaneous replacements and address two optimization criteria: cost-rate minimization and availability maximization.

Let the random variable Y_r denote the duration of reactive replacement with finite mean β_r. Similarly, let the random variable Y_p denote the the duration of preventive replacement with finite mean β_p, where $\beta_r \geq \beta_p \geq 0$. We assume that Y_r and Y_p are independent of the replacement age T. We denote the long-run expected cost-rate function under non-instantaneous replacements as $\mathcal{C}(T)$, and its optimal solution as $T^*_{\mathcal{C}}$. Then, the optimization problem for minimizing cost-rate under non-instantaneous replacements is:

$$\min_{T>0} \mathcal{C}(T) = \frac{\int_0^T c_r(x) f(x) dx + c_p(T)\bar{F}(T)}{\int_0^T \bar{F}(x) dx + \beta_r F(T) + \beta_p \bar{F}(T)}. \tag{3.4}$$

Hence, $\Omega(T)$ is a special case of $\mathcal{C}(T)$ as they coincide when $\beta_r = \beta_p = 0$. For *constant replacement costs*, [87] provide analytical conditions under which $\mathcal{C}(T)$ is monotone or quasi-convex. Here, we study the properties of function $\mathcal{C}(T)$ for *age-dependent* replacement costs.

In problem (3.4), we do not incorporate downtime cost. Instead, we are interested in comparing the optimal preventive replacement policies under two optimization criteria, namely, cost-rate minimization and availability maximization. Thereby, we can explore the degree to which age-dependent replacement costs may exacerbate the difference in the cost-rate minimizing and availability maximizing policies.

Let the system availability function and its optimal solution be denoted by $\mathcal{A}(T)$ and $T^*_{\mathcal{A}}$, respectively. Then, the optimization problem for maximizing availability is:

$$\max_{T>0} \mathcal{A}(T) = \frac{\int_0^T \bar{F}(x) dx}{\int_0^T \bar{F}(x) dx + \beta_r F(T) + \beta_p \bar{F}(T)}; \tag{3.5}$$

see, e.g., [96] for the derivation of (3.5). Appendix A provides a discussion on the properties of function $\mathcal{A}(T)$ and the existence and uniqueness of $T^*_{\mathcal{A}}$. The function $\mathcal{A}(T)$ is commonly referred to as steady-state availability [114], or limiting interval availability [96], or availability [87]. Throughout this chapter, we use the term availability. For some studies on the availability of deteriorating systems we refer the reader to [87, 96, 138, 35].

In Section 3.3, we study the properties of a special case of the cost-rate function $\mathcal{C}(T)$, namely that in which $\beta_p = \beta_r = 0$ (instantaneous replacements), and show that under mild assumptions, the cost-rate is quasi-convex with a unique finite optimal solution (see Theorem 6). Next, Theorem 9 establishes a similar result for cases in which $\beta_p = \beta_r > 0$.

Theorem 9. *Let $\beta_p = \beta_r > 0$. Under Assumption 6, $\mathcal{C}(T)$ is quasi-convex, there exists a unique finite solution $T^*_{\mathcal{C}}$ to problem* (3.4)*, and $\mathcal{C}(T^*_{\mathcal{C}}) = \big(c_r(T^*_{\mathcal{C}}) - c_p(T^*_{\mathcal{C}})\big)h(T^*_{\mathcal{C}}) + c_p'(T^*_{\mathcal{C}})$. Also, $T^*_{\mathcal{C}} < T^*_{\mathcal{A}} = +\infty$.*

The condition of Theorem 9, i.e., $\beta_p = \beta_r$, implies that the expected durations of reactive and preventive replacement are equal. This scenario arises in applications where the amount of time required to replace or perfectly repair a system is independent of whether the system is failed or still in a working condition. The result of Theorem 9 indicates that for this special case, the properties of $\mathcal{C}(T)$ are similar to those of $\Omega(T)$ (i.e., the cost-rate function under instantaneous replacements), and that the optimal cost-rate minimizing policy is unique and finite. Moreover, under the availability maximization criterion, there is no benefit in performing preventive maintenance; that is, $T^*_{\mathcal{A}} = +\infty$. Hence, the optimal policies under cost-rate minimization and availability maximization are significantly different.

Under age-dependent replacement costs, when $\beta_p < \beta_r$, it is difficult to analytically obtain easily verifiable conditions under which $\mathcal{C}(T)$ has a certain form. This difficulty arises from the fact that the properties of function $\mathcal{C}(T)$ depend on the combination of all the parameters involved; that is, the time to failure distribution, replacement durations, and replacement cost functions. Hence, we resort to exploring the properties of $\mathcal{C}(T)$ numerically. Our numerical experiments suggest that when replacement durations are considerably smaller than the expected time to failure of the system (μ), the cost-rate function is generally quasi-convex; thus, the optimal solution $T^*_{\mathcal{C}}$ is unique and finite. For this scenario, Example 11 depicts the cost-rate function $\mathcal{C}(T)$ under different forms of replacement cost functions.

Example 11 (Case of $\beta_p < \beta_r \ll \mu$). Let X follow a Weibull distribution with shape and scale parameter values 4 and 10, respectively. Hence, $\mu = 9.06$. Moreover, let $\beta_p = 0.2$, $\beta_r = 0.6$ (note that we have $\beta_p < \beta_r \ll \mu$), $c_p(t) = 4 + g(t)$ and $c_r(t) = 15 + g(t)$. Figure 10 depicts the cost-rate function $\mathcal{C}(T)$ vs. T for $g(t) \in \{0, \sqrt{t}, t, t^2\}$. ■

In Figure 10, we impose the conditions of Assumption 6 to avoid scenarios in which the replacement cost functions are step-like and result in multiple local minima (recall Figure 7). Hence, the cost-rate function $\mathcal{C}(T)$ is quasi-convex and the cost-rate minimizing policy is unique and finite. Based on our numerical study, these conclusions consistently hold across a wide range of functional forms of replacement costs and time to failure distributions.

For the scenarios depicted in Figure 10, Table 6 illustrates (*i*) the optimal cost-rate minimizing replacement times; (*ii*) the increase in cost-rate induced by never performing preventive maintenance; (*iii*) the increase in cost-rate induced by adopting the availability maximizing policy; (iv) and the decrease in availability induced by adopting the cost-rate minimizing policy. When replacement costs are constant (i.e., scenario (a) in Table 6), [87] show that if the ratio of reactive to preventive replacement costs is larger than that of replacement durations, then the optimal cost-rate minimizing policy replaces earlier than its availability maximizing counterpart; and that if the two ratios are equal, then the two policies coincide. Hence, $T_{\mathcal{C}}^*$ under scenario (a) is relatively close to $T_{\mathcal{A}}^* = 6.42$, and adopting a cost-rate minimizing or availability maximizing policy performs well under both optimization criteria. A similar conclusion holds in scenarios (b) and (c); however, it is violated in scenario (d) where the replacement costs increase rapidly in system age.

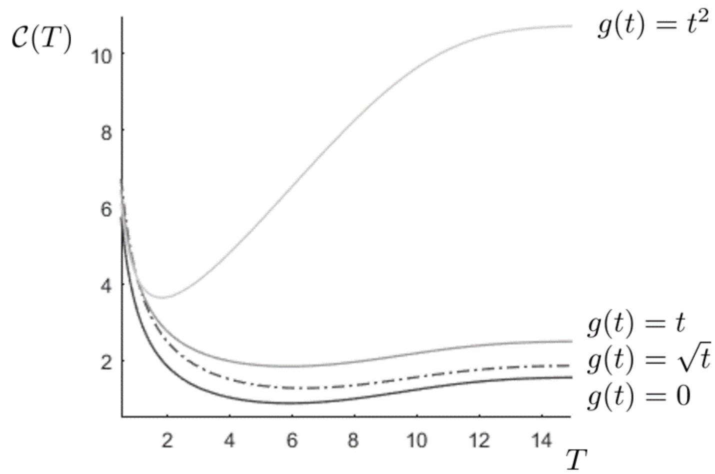


Figure 10: Case of $\beta_p < \beta_r \ll \mu$. Cost-rate function under different functional forms of replacement costs and the parameter values of Example 11 (we assume $c_p(t) = 4 + g(t)$ and $c_r(t) = 15 + g(t)$). Because replacement durations are considerably smaller than the expected time to failure, the cost-rate function is quasi-convex in all scenarios, and the optimal solution to (3.4) is unique and finite (i.e., it is optimal to perform preventive maintenance).

Next, we again assume that $\beta_p < \beta_r$, but that these values are not significantly smaller than μ. For this scenario, Example 12 depicts the cost-rate function $\mathcal{C}(T)$ under different forms of the replacement cost functions.

Example 12 (Case of $\beta_p < \beta_r \leq \mu$). Let X follow a Weibull distribution with shape and scale parameter values 4 and 10, respectively. Hence, $\mu = 9.06$. Moreover, we let $\beta_p = 3$, $\beta_r = 9$,

Table 6: Case of $\beta_p < \beta_r \ll \mu$. (i) Cost-rate minimizing replacement age, (ii) percent increase in cost-rate induced by never performing preventive maintenance (i.e., $T = +\infty$), (iii) percent increase in cost-rate induced by setting $T = T^*_{\mathcal{A}} = 6.42$, and (iv) percent increase in availability induced by setting $T = T^*_{\mathcal{C}}$ under different age-dependent replacement costs and the parameter values in Example 11.

		(i)	(ii)	(iii)	(iv)
Scenario	Replacement costs	$T^*_{\mathcal{C}}$	$\frac{\mathcal{C}(+\infty)-\mathcal{C}(T^*_{\mathcal{C}})}{\mathcal{C}(T^*_{\mathcal{C}})}$	$\frac{\mathcal{C}(T^*_{\mathcal{A}})-\mathcal{C}(T^*_{\mathcal{C}})}{\mathcal{C}(T^*_{\mathcal{C}})}$	$\frac{\mathcal{A}(T^*_{\mathcal{A}})-\mathcal{A}(T^*_{\mathcal{C}})}{\mathcal{A}(T^*_{\mathcal{A}})}$
(a)	$c_r(t) = 15$ $c_p(t) = 4$	5.90	76.7%	0.05%	0.03%
(b)	$c_r(t) = 15 + \sqrt{t}$ $c_p(t) = 4 + \sqrt{t}$	6.34	46.32%	0.02%	≈0
(c)	$c_r(t) = 15 + t$ $c_p(t) = 4 + t$	5.88	35.45%	0.47%	0.03%
(d)	$c_r(t) = 15 + t^2$ $c_p(t) = 4 + t^2$	1.80	195.42%	90.42%	6.21%

$c_p(t) = 4 + g_p(t)$ and $c_r(t) = 10 + g_r(t)$. Figure 11 depicts the cost-rate function $\mathcal{C}(T)$ vs. T for different functional forms of $g_p(t)$ and $g_r(t)$. ■

In Figure 11, it is apparent that the cost-rate function is no longer quasi-convex in general (except when $g_r(t) = g_p(t) = 0$ or $g_r(t) = g_p(t) = t$), and may have multiple local optima. In fact, the form of $\mathcal{C}(T)$ and the value of the cost-rate minimizing policy is very sensitive to the parameter values and replacement cost functions. For example, when $g_r(t)$ and $g_p(t)$ are approximately linear functions with a slope of 1, the optimal cost-rate minimizing policy is to never perform preventive maintenance; otherwise, it is finite. We also want to note that in our numerical study, a similar conclusion does not necessarily hold for cases where $g_r(t)$ and $g_p(t)$ are linear; rather, it depends on the combination of the other parameter values. For example, when $\beta_r = 9$, $\beta_p = 1$, $c_r(t) = 90 + t^2$ and $c_p(t) = 40 + t^2$, the optimal policy is to never perform preventive maintenance, even though

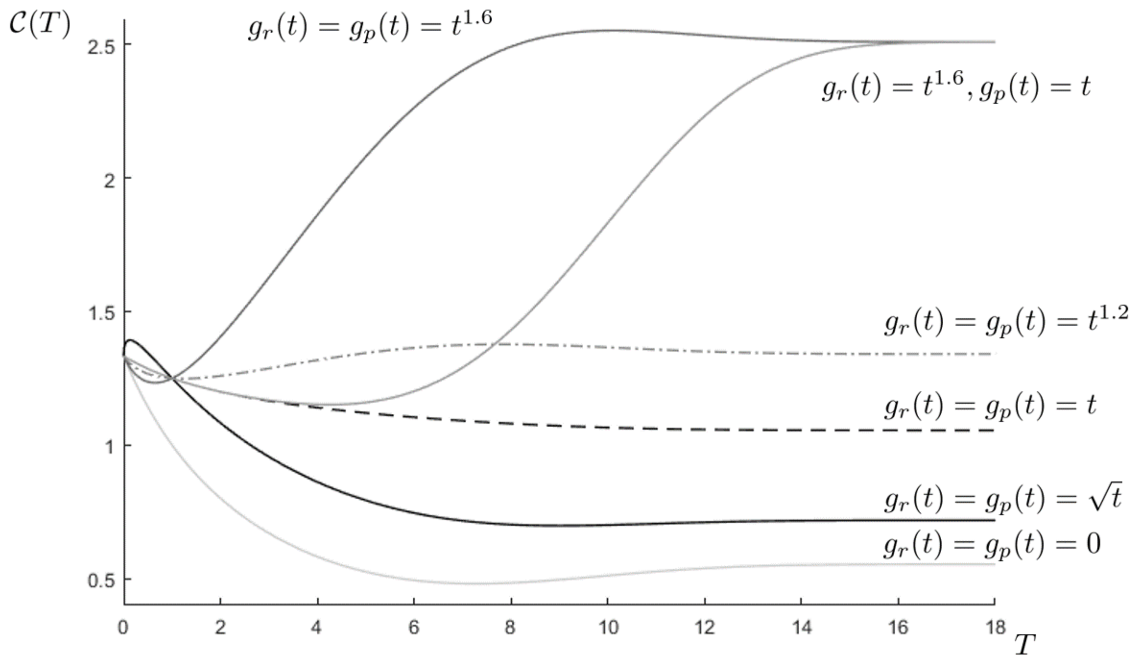


Figure 11: Case of $\beta_p < \beta_r \leq \mu$. Cost-rate function under different forms of replacement cost functions and the parameter values of Example 12 (we assume $c_r(t) = 10 + g_r(t)$ and $c_p(t) = 4 + g_p(t)$). Because the replacement durations are only slightly smaller than the expected time to failure, the properties of $\mathcal{C}(T)$ are very sensitive to the replacement cost functions. Moreover, when $g_r(t) = g_p(t) \approx t$, it is optimal to never perform preventive maintenance.

the reactive replacement cost increases rapidly. Figure 11 indicates that misspecification of the replacement cost functions can increase the cost-rate significantly.

Table 7 is the counterpart of Table 6 for the scenarios depicted in Figure 11. Because the optimal replacement age is sensitive to the forms of the replacement cost functions, the decrease in system availability is significant if the cost-rate minimizing policy is adopted, and vice versa. Hence, when replacement durations are not considerably smaller than the expected time to failure, maintenance planners should (i) be concerned about accurately modeling the age-dependent replacement costs, and (ii) identify priorities in the optimization criteria because adopting a replacement policy under one criterion may significantly compromise the other criterion. In these scenarios, obtaining Pareto efficient solutions may be of interest.

Table 7: Case of $\beta_p < \beta_r \leq \mu$. (i) Cost-rate minimizing replacement age, (ii) percent increase in cost-rate induced by never performing preventive maintenance (i.e., $T = +\infty$), (iii) percent increase in cost-rate induced by setting $T = T^*_{\mathcal{A}} = 6.42$, and (iv) percent increase in availability induced by setting $T = T^*_{\mathcal{C}}$ under different age-dependent replacement costs and the parameter values in Example 12.

		(i)	(ii)	(iii)	(iv)
Scenario	Replacement costs	$T^*_{\mathcal{C}}$	$\frac{\mathcal{C}(+\infty)-\mathcal{C}(T^*_{\mathcal{C}})}{\mathcal{C}(T^*_{\mathcal{C}})}$	$\frac{\mathcal{C}(T^*_{\mathcal{A}})-\mathcal{C}(T^*_{\mathcal{C}})}{\mathcal{C}(T^*_{\mathcal{C}})}$	$\frac{\mathcal{A}(T^*_{\mathcal{A}})-\mathcal{A}(T^*_{\mathcal{C}})}{\mathcal{A}(T^*_{\mathcal{A}})}$
(a)	$c_r(t) = 10$ $c_p(t) = 4$	7.28	15.09%	1.15%	0.84%
(b)	$c_r(t) = 10 + \sqrt{t}$ $c_p(t) = 4 + \sqrt{t}$	9.02	2.88%	4.79%	18.01%
(c)	$c_r(t) = 10 + t$ $c_p(t) = 4 + t$	$+\infty$	0	4.10%	18.02%
(d)	$c_r(t) = 10 + t^{1.2}$ $c_p(t) = 4 + t^{1.2}$	1.22	7.39%	9.69%	52.64%
(e)	$c_r(t) = 10 + t^{1.6}$ $c_p(t) = 4 + t$	4.24	117.91%	7.02%	7.00%
(f)	$c_r(t) = 10 + t^{1.6}$ $c_p(t) = 4 + t^{1.6}$	0.65	103.42%	88.66%	70.98%

Finally, Table 8 summarizes our findings on the properties of functions $\mathcal{C}(T)$ and $\mathcal{A}(T)$ and optimization problems (3.4) and (3.5). Note that, we do not study the case where replacement durations are larger than the expected time to failure as such scenarios are uncommon in practice.

Table 8: Summary of findings on the functions $\mathcal{C}(T)$ and $\mathcal{A}(T)$, and the cost-rate minimizing and availability maximizing policies, under different replacement duration scenarios.

	$\min_{T>0} \mathcal{C}(T)$	$\max_{T>0} \mathcal{A}(T)$
$\beta_p = \beta_r = 0$	Under mild conditions, $\mathcal{C}(T)$ is quasi-convex, and there exists a unique and finite optimal solution (Theorem 6)	$\mathcal{A}(T) = 1$ for all T
$\beta_p = \beta_r > 0$	Under mild conditions, $\mathcal{C}(T)$ is quasi-convex, and there exists a unique and finite optimal solution (Theorem 9)	$\mathcal{A}(T)$ is monotone increasing, and the optimal solution is infinite
$\beta_p < \beta_r \ll \mu$	$\mathcal{C}(T)$ is generally quasi-convex, and there exists a unique and finite optimal solution (Example 11)	$\mathcal{A}(T)$ is quasi-convex, and there exists a unique and finite optimal solution (Appendix A)
$\beta_p < \beta_r \leq \mu$	The properties of $\mathcal{C}(T)$ are sensitive to the functional form of replacement costs, and the optimal solution may be finite or infinite (Example 12)	$\mathcal{A}(T)$ is quasi-convex, and there exists a unique and finite optimal solution (Appendix A)

3.6 Concluding Remarks

We consider age-replacement policies (without minimal repair) for systems with increasing replacement costs under both instantaneous and non-instantaneous replacement durations. When replacements are instantaneous, we provide sufficient conditions under which there exists a unique optimal solution that minimizes the long-run expected cost-rate. Moreover, we examine how age-dependent replacement costs affect the optimal replacement time and the optimal cost-rate compared to those under constant replacement costs. We also explore the increase in long-run expected cost-rate when replacement costs are age-dependent but a replacement policy based on constant replacement costs is implemented. Under convex replacement costs, the increase in cost-rate is large if the policy is based on replacement costs that have a small reactive to preventive cost ratio. The opposite holds when replacement costs increase slowly and their ratio decreases in age.

Under non-instantaneous replacements, we study the form of the cost-rate function and compare optimal cost-rate minimizing and availability maximizing replacement policies. Our experiments indicate that ignoring age-dependency can result in both underestimating the long-run expected cost-rate and overestimating system availability. We also learn that when replacement durations are comparable to the expected time to failure, the form of the cost-rate function and its optimal solution are sensitive to the replacement cost functions. Hence, it is important that maintenance planners accurately model the age-dependent replacement costs.

4.0 Optimal Condition-Based Maintenance via a Mobile Maintenance Resource

4.1 Introduction

In this chapter, we consider a condition-based maintenance setting where a set of identical, geographically distributed assets degrade stochastically over time and may fail without proper intervention. Failures do not mandate immediate maintenance, but do incur downtime costs. A single maintenance resource is tasked with traveling among and maintaining these assets. We model the asset deterioration process as a completely observable discrete-time Markov chain, and the location of the maintenance resource and assets using a graph (network). In this graph, nodes represent possible geographical locations for the maintenance resource, which include both auxiliary and asset locations. Traversing the auxiliary locations allows the maintenance resource to reach asset locations. The maintenance resource may also idle at the auxiliary or asset locations at any time. Edges represent links between nodes along which the maintenance resource may travel. We assume that assets are not co-located, and hence, the maintenance resource travels between asset nodes to carry out repairs, which restore the assets to as-good-as-new.

We seek to (i) dynamically obtain the optimal actions (repair, reposition, idle) for the maintenance resource as a function of the conditions of the assets and the current location of the resource to minimize total expected discounted costs which include downtime, travel, and maintenance expenses; (ii) study the structural properties of the optimal policies; (iii) provide insights on how current and future locations of the resource can be exploited to perform proximal maintenance, and how the maintenance resource can be strategically repositioned in anticipation of maintenance needs; (iv) explore how graph structure affects positioning and maintenance decisions; and (v) design implementation-friendly heuristic policies.

The majority of studies on condition-based maintenance focus on adaptively determining when to perform maintenance as a function of the condition of a single asset or multiple co-located assets with economic dependencies; we refer the reader to [6,] for a survey. To the best of our knowledge, the only existing work on condition-based maintenance for dispersed assets is that in [62]. In their work, the assets are dispersed on a cycle graph and are visited in a fixed order by a maintenance resource, hence no decisions are made on how to dynamically reposition the resource. Our modeling framework, on the other hand, generalizes the graph configuration and addresses both the timing of maintenance interventions and the dynamic repositioning of the maintenance resource. A recent survey points out that condition-based maintenance of geographically dispersed assets is an open research direction [45].

Other related studies consider preventive maintenance for geographically dispersed assets, but in the context of *time-based maintenance*. These studies include [51, 60, 99, 85, 107, 117] who consider *deterministic* settings where a set of maintenance tasks are predefined or scheduled based on the assets' ages and are carried out by multiple resources. Similarly, the studies in [30, 31] determine optimal routes *a priori* that are updated once an asset fails. The studies in [38, 72], asset locations are modeled in a flow network, where asset failures on graph nodes can interrupt the flow of material in the network; but again, lifetime distributions are used to make maintenance decisions *a priori*. Other studies consider routing mobile assets to fixed maintenance facilities in which maintenance schedules are predefined or determined based on number of operating hours [55, 112, 123]. Unlike this body of literature, in our problem, maintenance actions and movements of the maintenance resource are dynamically prescribed in response to changes in the assets' conditions, enabling maintenance planners to exploit the cost-saving benefits of condition monitoring.

Our research also has loose connections with the dynamic traveling repairman problem [20, 21, 22], the medical unit dispatching literature [5, 73, 90, 91], and dynamic repositioning problems for general applications [18, 19, 49]. The author in [125] studies optimal waiting strategies when anticipating service requests from multiple customer locations. Other studies consider dynamic vehicle routing problems in which some of the customer demands are known *a priori* while others are dynamic [17, 79, 80]. In these studies, *demands* arrive stochastically (either in some Euclidean space or a general network), and can be viewed in a maintenance context as failed assets that (unlike our assets) require immediate maintenance. Our work, however, focuses on applications with condition monitoring technologies put in place to enable maintenance planners to maintain assets before failure. Hence, the models in this body of literature cannot adequately address the trade-offs in our maintenance setting.

Finally, another somewhat relevant body of literature is that on the inventory routing problem, in which a supplier decides each period how much to deliver to each customer and how to assign trucks/routes (see [42] for a review). In these studies inventory is often distributed from a depot to a set of customers using a fleet of vehicles with limited capacity [32, 75, 76] or is redistributed between different stations as in a bike sharing system [27, 28]. In relation to our problem, customer (or station) inventory levels can be viewed as asset deterioration conditions, and replenishment decisions as maintenance interventions. That said, in inventory routing, demand is assumed to be either deterministic or independently and identically distributed at each customer; in our setting, the probability of transitioning to a new deterioration condition depends on the current deterioration condition. Our model also allows for transitioning to better conditions even in the absence of maintenance. Moreover, similar to the vehicle routing literature, these studies only identify delivery routes that must start and end at a depot and do not allow idling. Lastly, the cost structures in these problems differ from ours in that they may include inventory holding costs, delivery rewards, or shortage penalties.

Table 9 summarizes the most relevant literature and highlights our contributions. The first column lists the attributes of interest. For example, the second attribute labeled as "Allow maintenance before failure" signifies whether maintenance is allowed at failure only or if maintenance is allowed before failure as well. Similarly, the third and fourth attributes determine whether maintenance or positioning decisions are dynamic or pre-planned. The last attribute indicates the methodology used (MDP and QT stand for Markov decision processes and Queueing Theory, respectively). Notice that our work is the first to jointly address optimal condition-based maintenance and dynamic positioning of a maintenance resource. Table 9 also identifies potential future directions, for instance, considering condition-based maintenance for dispersed assets with partially observable conditions.

The remainder of the chapter is structured as follows. Section 4.2 formulates a Markov decision process model to obtain the optimal actions of the maintenance resource. Section 4.3 establishes structural properties of the optimal policy obtained from our theoretical derivations. Section 4.4 provides insights on the structure of the optimal policy obtained from

Table 9: Summary of the related literature and our contributions.

	Our model	[125]	[18, 19]	[5, 73, 90, 91]	[62]	[2]	[21, 22]	[70]	[79]	[80]	[17]	[51, 69, 85, 107, 117]	[38, 99]	[30, 31]	[72]	[60]
Condition-based maintenance	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Allow maintenance before failure	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓
Dynamic maintenance decisions	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗
Dynamic positioning decisions	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗
Allow idling	✓	✓	✗	✓	✗	✓	✗	✓	✓	✓	✗	✗	✗	✗	✗	✓
Partially observable conditions	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
Multiple maintenance resources	✗	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗	✓	✗	✗	✓	✓
Random service times	✗	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✗
Random travel times	✗	✗	✗	✓	✗	✓	✓	✓	✓	✗	✓	✗	✗	✗	✗	✗
Methodology	MDP					QT			Deterministic Optimization							

numerical observations. In particular, we explore scenarios in which maintenance is earlier compared to a single-asset setting and introduce the concept of *proximal maintenance*. Moreover, we explore use of graph centrality measures to identify promising idling locations for the maintenance resource and present insights on optimal repositioning decisions. Sections 4.5 analyzes the sensitivity of several metrics of interest to various parameter values (e.g., costs or transition probabilities). Section 4.6 provides easy-to-implement heuristic policies and analyzes their performance. Finally, Section 4.7 summarizes our findings. Appendices C - F contain the proofs for all established results and additional numerical examples.

4.2 Model Formulation

Consider a single maintenance resource responsible for maintaining a set of n_M identical assets. We use graph $G = (V, E)$ to capture the geographical locations of the assets and possible repositioning movements between them executed by the maintenance resource. The set of nodes V consists of two disjoint subsets V_M and V_T, i.e., $V = V_M \cup V_T$ and $V_M \cap V_T = \emptyset$. Nodes in $V_M = \{1, \ldots, n_M\}$ represent the locations of the assets. Nodes in $V_T = \{l_1, \ldots, l_{n_T}\}$ are "auxiliary" nodes that are used to model travel between the assets. The maintenance resource idles in or travels through these locations to reach asset locations. Hence, $V = \{1, \ldots, n_M, l_1, \ldots, l_{n_T}\}$. In our infinite horizon model, time is discretized and it is assumed that the maintenance resource can traverse one graph edge per time period. For each pair of nodes $b, b' \in V$, $b \neq b'$, the edge (b, b') is contained in the set of edges E if and only if it is possible to move from node b to node b' within one time period. The maintenance resource can traverse the edges of the graph and may repair an asset once it is located in an asset node. Similar to traversing an edge, a repair action is assumed to take one time period and restores the asset to as-good-as-new. We assume that the maintenance resource has no home-base requirements and is available at all decision epochs.

The assets are prone to deterioration over time, and their deterioration condition is remotely monitored and fully observed by sensors. The assets are assumed to deteriorate independently according to a discrete-time Markov chain. The deterioration conditions are denoted by $0, 1, \ldots, \Delta - 1, \Delta$; where 0 represents the as-good-as-new condition, and $\Delta < \infty$ represents the failed condition in which the asset is "down." Let $\mathcal{K} = \{0, 1, \ldots, \Delta - 1, \Delta\}$. We denote the transition probability matrix for the discrete-time Markov chain by P, with elements $P_{i,j}$ denoting the probability of transitioning to deterioration condition j from i in one time period.

Figure 12 depicts an example of a graph with four assets ($n_M = 4$) and four possible deterioration conditions ($\Delta = 3$), as well as five auxiliary nodes ($n_T = 5$). Auxiliary nodes capture travel distances and possible travel routes between assets. For instance, it takes six units of time to travel from Asset 1 to Asset 4. That is, the maintenance resource must traverse the auxiliary nodes $l_1, l_2, ..., l_5$ to reach Asset 4 when repositioning from Asset 1.

As depicted in Figure 12, we generally consider settings where travel durations between assets are relatively longer than repair times (recall that a repair action takes one unit of time). Our model is motivated by applications that arise in fulfilment centers, warehouses and manufacturing plants, data centers, as well as recently developed satellite maintenance systems [10, 26, 110]. In these applications, maintenance tasks often include simple and quick fixes or component replacement. For instance in a data center, maintenance tasks include replacing generators, switches, and backup batteries [142]. Thus, a human technician or a servicing robot travels long distances between servers and performs quick repairs upon arrival.

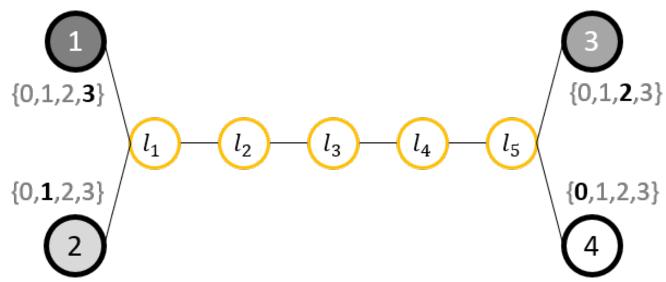


Figure 12: Example of a (2,4)-banana graph [119] with four asset nodes ($V_M = \{1, 2, 3, 4\}$) and five auxiliary nodes ($V_T = \{l_1, \dots, l_5\}$). Each asset can be in one of four deterioration conditions ($\mathcal{K} = \{0, 1, 2, \Delta = 3\}$); darker asset nodes indicate worse conditions. Asset conditions are also indicated on the labels next to the assets.

Next, we formulate the components of our MDP model.

State Space. The state of the MDP includes the deterioration conditions of the assets and the location of the maintenance resource, as these are the only pieces of information needed to determine the resource's next action. Specifically, let x_i denote the deterioration condition of asset $i \in V_M$ and $\boldsymbol{x} = (x_1, \dots, x_{n_M})$ be the vector of deterioration conditions of all the assets. Furthermore, let $l \in V$ denote the current location of the maintenance resource. The state of the MDP is then $\boldsymbol{s} = (\boldsymbol{x}, l)$, and the state space S is

$$S = \Big\{(\boldsymbol{x}, l) : \boldsymbol{x} \in \mathcal{K}^{n_M}, l \in V\Big\}.$$

Actions. We consider three types of actions. When the maintenance resource is in an asset location, a repair may be carried out (action R); the maintenance resource may also travel to an adjacent node b (action T_b); or, the maintenance resource can do nothing, i.e., idle (action DN). When the maintenance resource is in an auxiliary node, it may either travel to an adjacent node or do nothing. The set of allowable actions, denoted by $A_{\boldsymbol{s}}$, depends on the current state $\boldsymbol{s} = (\boldsymbol{x}, l)$ and is expressed as

$$A_{\boldsymbol{s}} = \begin{cases} \{R, DN\} \bigcup \big(\bigcup\limits_{b:(l,b)\in E} \{T_b\} \big), & \text{if } l \in V_M, \\ \{DN\} \bigcup \big(\bigcup\limits_{b:(l,b)\in E} \{T_b\} \big), & \text{if } l \in V_T. \end{cases}$$

Transition Probabilities. Let $p(\boldsymbol{s}'|\boldsymbol{s}, a)$ denote the probability of transitioning to state $\boldsymbol{s}' = (\boldsymbol{x}', l')$ when the current state is $\boldsymbol{s}$ and action a is chosen. Recall that repair actions are perfect, i.e., restore the asset to as-good-as-new condition. If the repair action is chosen (i.e., $a = R$), then the only possible transitions are to states in which the repaired asset is as-good-as-new and maintenance resource location l' is the same as the current location l:

$$p(\boldsymbol{s}'|\boldsymbol{s}, R) = \begin{cases} \prod_{j\in V_M\setminus\{l\}} P_{x_j, x'_j}, & \text{if } x'_l = 0, l' = l; \\ 0, & \text{otherwise.} \end{cases}$$

On the other hand, if the do nothing action is chosen (i.e., $a = DN$), then the only possible transitions are to states in which maintenance resource location l' is the same as the current location l:

$$p(\boldsymbol{s}'|\boldsymbol{s}, DN) = \begin{cases} \prod_{j\in V_M} P_{x_j, x'_j}, & \text{if } l' = l; \\ 0, & \text{otherwise.} \end{cases}$$

Finally, if travel action to node b is chosen (i.e., $a = T_b$), then the only possible transitions are to states in which maintenance resource location l' is b:

$$p(\boldsymbol{s}'|\boldsymbol{s}, T_b) = \begin{cases} \prod_{j\in V_M} P_{x_j, x'_j}, & \text{if } l' = b; \\ 0, & \text{otherwise.} \end{cases}$$

Rewards. Three types of costs may be incurred: repair, downtime, and travel. The cost of repairing an asset in condition $k \in \mathcal{K}$ is denoted by $c_R(k) \geq 0$. We assume that the repair cost is non-decreasing in deterioration condition because it may be more costly to repair or replace a highly deteriorated asset compared to a healthier asset. That is, $0 \leq c_R(0) \leq \ldots \leq c_R(\Delta)$. This assumption is common in the maintenance optimization literature [56, 15]. A per unit downtime cost $c_D \geq 0$ is incurred in each period by any asset that is not functioning either because it is in the failed state or is undergoing repair. A travel cost $c_T \geq 0$ is incurred for traversing any edge $(b, b') \in E$.

Using the above notation, the state and action dependent immediate costs $r(\boldsymbol{s}, a)$ can be expressed as

$$r(\boldsymbol{s}, R) = c_R(x_l) + c_D + \sum_{j \in V_M \setminus \{l\}} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}}, \tag{4.1}$$

$$r(\boldsymbol{s}, DN) = \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}}, \tag{4.2}$$

$$r(\boldsymbol{s}, T_b) = c_T + \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}}. \tag{4.3}$$

Equation (4.1), for example, can be interpreted as follows. If the repair action is chosen in state $\boldsymbol{s} = \big((x_1, \ldots, x_l, \ldots, x_{n_M}), l\big)$, then the immediate cost is the summation of asset l's repair and downtime cost, plus the downtime cost of any other failed assets. Equations (4.2) and (4.3) can be interpreted in a similar manner.

Value Function. The overall goal is to minimize the long-run total expected discounted cost by choosing optimal actions as a function of the MDP state. Let $v_{\boldsymbol{s}}$ be the expected minimal discounted cost-to-go starting from state $\boldsymbol{s} = (\boldsymbol{x}, l)$ and let $\lambda \in [0, 1)$ be a discount factor. Then

$$v(\boldsymbol{s}) = v(\boldsymbol{x}, l) = \begin{cases} \min\Big\{\mathcal{R}(\boldsymbol{s}), \mathcal{DN}(\boldsymbol{s}), \min\limits_{b:(l,b) \in E} \mathcal{T}_b(\boldsymbol{s})\Big\}, & \text{if } l \in V_M, \\ \min\Big\{\mathcal{DN}(\boldsymbol{s}), \min\limits_{b:(l,b) \in E} \mathcal{T}_b(\boldsymbol{s})\Big\}, & \text{if } l \in V_T, \end{cases} \tag{4.4}$$

where

$$\mathcal{R}(\boldsymbol{s}) \equiv c_R(x_l) + c_D + \sum_{j \in V_M \setminus \{l\}} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}' \in \mathcal{K}^{n_M} \\ \text{s.t. } x'_l = 0}} \prod_{j \in V_M \setminus \{l\}} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', l), \tag{4.5}$$

$$\mathcal{DN}(\boldsymbol{s}) \equiv \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', l), \text{ and} \tag{4.6}$$

$$\mathcal{T}_b(\boldsymbol{s}) \equiv c_T + \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', b). \tag{4.7}$$

Equations (4.5)-(4.7) represent the total expected discounted cost-to-go starting from state $\boldsymbol{s}$ and choosing action repair, do nothing, and travel to node b, respectively. The optimal action in state $\boldsymbol{s}$, denoted by $a^*(\boldsymbol{s})$, is the one that obtains the minimum on the right-hand side of equation (4.4). In the remainder of the chapter we use the value iteration algorithm to compute the value function in (4.4) and obtain the optimal actions [106].

4.3 Structural Properties

For analyzing the structural properties of the value function and optimal policies, we first provide the definition of an increasing failure rate (IFR) stochastic matrix [13].

Definition 1. Let $\sum_{j=k}^{\Delta} P_{i,j}$ be nondecreasing in i for all $k \in \mathcal{K}$. Then, matrix P has the increasing failure rate (IFR) property.

The IFR property indicates that assets deteriorate faster in worse conditions (see e.g., [1, 65]). The deterioration processes of many real world applications exhibit both IFR and upper-triangular properties [1, 29, 65, 67]. Upper-triangularity implies that asset condition cannot improve in the absence of maintenance interventions, but it is neither required for nor implied by the IFR property.

We next show that under the IFR property, the optimal value function is monotonically nondecreasing in each asset's condition when all other state variables remain fixed. All proofs are provided in Appendix A.

Proposition 10. *If P has the IFR property, then $v\big((x_1, \dots, x_i, \dots, x_{n_M}), l\big)$ is nondecreasing in x_i for fixed l and x_j, $j \neq i$.*

Using Proposition 10, we establish sufficient conditions for the existence of an optimal control-limit for the repair action with respect to the condition of the asset that is located in the position of the maintenance resource.

Theorem 10. *Consider the following two sets of conditions: (i) P has the IFR property and $c_R(k)$ is constant for all $k \in \mathcal{K}$; (ii) P has the IFR property, $c_R(k)$ is constant for all $k \in \mathcal{K} \setminus \{\Delta\}$ and $c_R(\Delta) - c_R(\Delta - 1) \leq c_D$. Under either set of conditions in (i) or (ii), if there exists a* *condition* x_i^* *such that* $a^*\big((x_1,\ldots,x_i^*,\ldots,x_{n_M}),i\big)=R$, *then* $a^*\big((x_1,\ldots,x_i,\ldots,x_{n_M}),i\big)=R$ *for all* $x_i\geq x_i^*$.

Theorem 10 implies that under mild conditions, when the maintenance resource is at an asset location, the optimal maintenance decision can be characterized by a repair threshold for that asset given the deterioration conditions of other assets. This control-limit structure is appealing because it can save computational effort and is easy to implement in practice [106]. In Section 4.6, we exploit this structure in developing heuristic policies.

The sufficient conditions of Theorem 10 ensure a control-limit structure by ruling out repair costs that are significantly higher in worse conditions; e.g., the condition $c_R(\Delta)-c_R(\Delta-1)\leq c_D$ ensures that the difference between the repair cost at failure and at $\Delta-1$ is bounded by the downtime cost. In scenarios in which repair costs are significantly higher in more deteriorated conditions compared to better conditions, the control-limit structure may be violated. That is, it may be optimal to repair an asset in a healthier condition, but suboptimal to repair it in a relatively more deteriorated condition. In the extreme of such instances, it may be optimal to abandon assets once they reach a certain level of deterioration. See examples in Appendix B. In many of our numerical examples we let the repair cost function take more general forms (e.g., monotone increasing) than those described in the conditions of Theorem 10. However, in these examples, the repair costs do not vary significantly between different deterioration conditions, and we observe that the optimal repair action follows a control-limit rule.

Next, using Proposition 10, Theorem 11 establishes conditions under which it is suboptimal to reposition to a location with a higher total expected discounted cost-to-go value.

Theorem 11. *Let* P *have the IFR and upper-triangular properties. Consider two adjacent locations* l *and* b, *i.e.,* $(l,b)\in E$. *If* $v(\boldsymbol{x},l)=\mathcal{T}_b(\boldsymbol{x},l)$, *then* $v(\boldsymbol{x},l)\geq\lambda v(\boldsymbol{x},b)$. *Moreover, consider the following two sets of conditions: (i)* $v(\boldsymbol{x},l)<\lambda v(\boldsymbol{x},b)$; *(ii)* $c_T>0$ *and* $v(\boldsymbol{x},l)\leq\lambda v(\boldsymbol{x},b)$. *If either (i) or (ii) holds, then* $v(\boldsymbol{x},l)<\mathcal{T}_b(\boldsymbol{x},l)$.

The first result in Theorem 11 demonstrates that if it is optimal to reposition to a particular location, then that location has a lower long-run expected discounted cost than the current location. The second result establishes the reverse case; that is, if the long-run expected discounted cost of a location is less than that of an adjacent location, then it is suboptimal to reposition to that adjacent location. When $v(\boldsymbol{x}, l) = \lambda v(\boldsymbol{x}, b)$ the result is violated only if the travel cost is zero.

A direct consequence of Theorem 11 is that the optimal action in a node with locally minimum value function cannot be traveling, and is instead idling or repairing. That is, if $v(\boldsymbol{x}, l) < \lambda \min_{b:(l,b)\in E} v(\boldsymbol{x}, b)$, then $v(\boldsymbol{x}, l) = \mathcal{DN}(\boldsymbol{x}, l)$ for $l \in V_T$ and $v(\boldsymbol{x}, l) = \min\{\mathcal{DN}(\boldsymbol{x}, l), \mathcal{R}(\boldsymbol{x}, l)\}$ for $l \in V_M$. Note that the upper-triangular property implies that asset conditions cannot improve in the absence of maintenance and thus, by Proposition 10 the value function at the adjacent location b cannot improve in the next decision epoch. Consequently, it is suboptimal to travel from node l to b in anticipation of an improvement in the value function.

4.4 Policy Insights

In this section, we discuss interesting insights on the structure of the optimal policy based on numerical experimentation. Specifically, first in Section 4.4.1, we discuss the factors that prompt "early" maintenance, that is, earlier than in a setting with a single asset. Then in Section 4.4.2, we characterize how the vector of deterioration conditions affects the optimal idling and repositioning decisions. Lastly in Section 4.4.3, we conduct a simulation study that identifies the locations in the graph that are most used for idling under the optimal policy and examine their relationship to graph structure. We build on these findings to design high-performance heuristic policies in Section 4.6.

4.4.1 Maintenance Thresholds

Recall that under the IFR property and the cost conditions in Theorem 10, maintenance decisions when in an asset location can be characterized by an optimal repair *threshold*. In the special case of a single asset with a dedicated maintenance resource, this optimal repair threshold only depends on the condition-dependent repair costs, downtime cost, and the asset's deterioration process. In our more general setting, however, the optimal thresholds are affected not only by these parameter values, but also by the conditions of the other assets, the relative distances between the assets, the underlying graph structure, and the current location of the maintenance resource. These novel features add to the complexity of the decision making process pertaining to the maintenance decisions.

In general, when the maintenance resource is at an asset location, it is often optimal to repair an asset earlier (i.e., in a less-deteriorated condition) than it would be if maintaining only that one asset in isolation; we refer to this phenomenon as *early* maintenance. That is, early maintenance implies that it is optimal to maintain an asset earlier in the multi-asset setting compared to the single-asset setting under the same costs and deterioration process.

We summarize our numerical observations in three important scenarios where *early* maintenance is optimal, namely, when (i) an asset is in a non-central and thus unfavorable location; (ii) multiple assets are deteriorated; and, (iii) the maintenance resource capitalizes on its proximity to an asset, which we refer to as *proximal maintenance*. Note that proximal maintenance is somewhat similar to opportunistic maintenance in that it exploits opportunities to save costs by maintaining early. However, opportunistic maintenance applies to multi-component systems and uses planned or unplanned downtime caused by one component to preventively maintain another component(s) [48, 134, 143]. Therefore, the events that trigger opportunistic and proximal maintenance are different.

These scenarios are presented in Examples 13-15 where we let $c_R(0) = 2$, $c_R(1) = 4$, $c_R(2) = 6$, $c_R(3) = 10$, $c_D = 10$, $c_T = 0.5$, and

$$P = \begin{bmatrix} 0.98 & 0.01 & 0.01 & 0 \\ 0 & 0.96 & 0.03 & 0.01 \\ 0 & 0 & 0.95 & 0.05 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

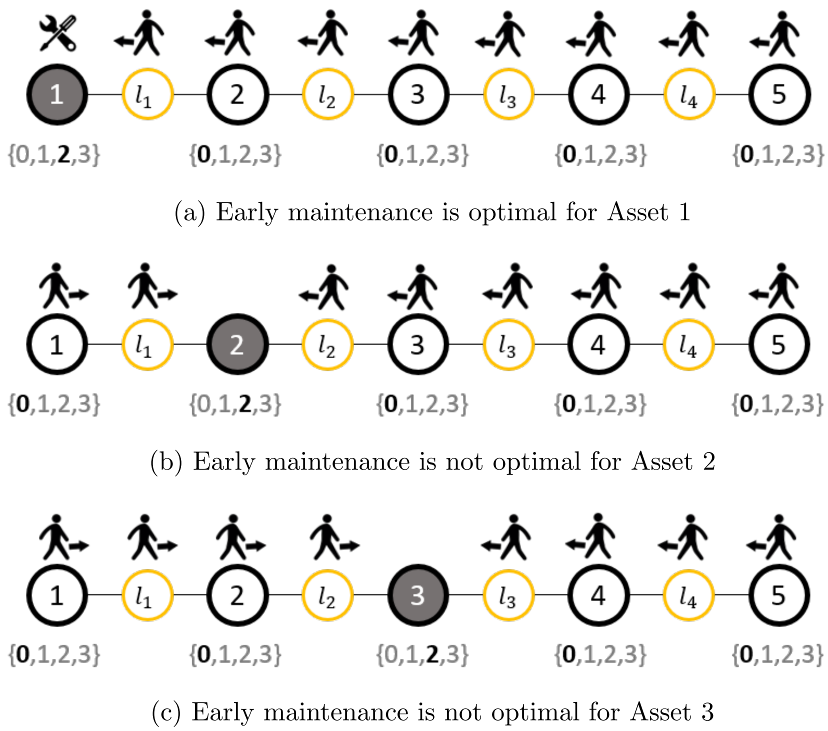
Under these parameter values, the optimal maintenance threshold for an asset in isolation is 3 (i.e., at failure). We obtain this value simply by solving the special case of our model with a single asset node.

Example 13 (Deteriorated Asset in a Non-Central Location). One scenario in which it is optimal to maintain an asset earlier than we would for that asset in isolation (i.e., earlier than deterioration level of 3 under these parameter values) arises when the maintenance resource is near a non-centrally located asset. Maintaining such assets early can be optimal because the maintenance resource can then reposition to more central locations; take Figure 13 as an example.

Figure 13 depicts three scenarios, each with one deteriorated asset. Note that in all three scenarios the deteriorated asset has the same deterioration level of 2, but early maintenance is only optimal in Figure 13a when the deteriorated asset is Asset 1, which is located in a relatively less-central location. Early intervention allows the maintenance resource to subsequently reposition to central nodes of the graph to possibly idle in anticipation of further deterioration of the assets. That is, although not presented in Figure 13, when all assets are as-good-as-new and the maintenance resource is at Asset 1, the optimal action is to travel toward l_1. ■

Example 14 (Multiple Assets Are Deteriorated). Another common scenario in which optimal early maintenance occurs arises when multiple assets are deteriorated. In such scenarios, by maintaining an asset early, the maintenance resource can subsequently reposition to the location of another deteriorated asset; take Figure 14 as an example. Comparing Figures 13c and 14 we observe that early intervention is optimal in the latter because another asset is deteriorated. ■

Example 15 (Proximal Maintenance). Early maintenance of an asset can also be optimal due to the proximity of the maintenance resource to that asset; we refer to this type of early maintenance as proximal maintenance. In such scenarios, it may not be optimal to travel towards a deteriorated asset either because other assets are more deteriorated, or because that asset is not sufficiently deteriorated to justify the costs associated with traveling toward that asset. However, if the maintenance resource is already at that asset, it may be optimal to perform early maintenance; see Figure 15 as an example where proximal maintenance is optimal for Asset 1. ■



(a) Early maintenance is optimal for Asset 1

(b) Early maintenance is not optimal for Asset 2

(c) Early maintenance is not optimal for Asset 3

Figure 13: An excerpt of the optimal policy for Example 13. Only in (a) is early maintenance optimal because the deteriorated asset is in a non-central location. That is, early maintenance of Asset 1 allows the maintenance resource to subsequently reposition to $\boldsymbol{l_1}$ which is more central to all assets. Icons represent optimal actions as follows: repair, travel in the indicated direction. The do nothing action is optimal in nodes with no icon.

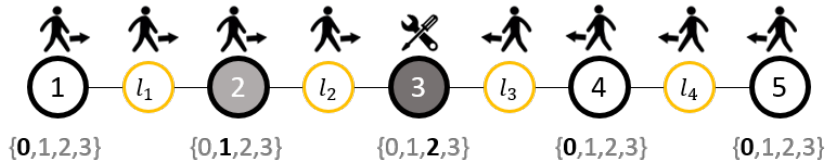


Figure 14: An excerpt of the optimal policy for Example 14. Early maintenance is optimal for Asset 3 because the maintenance resource can subsequently travel to Asset 2 in anticipation of its further deterioration.

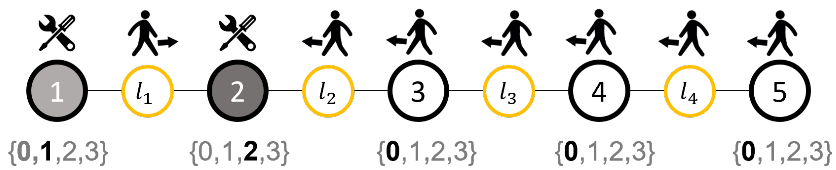


Figure 15: An excerpt of the optimal policy for Example 15. It is not optimal to travel towards Asset 1 in $\boldsymbol{l_1}$, however, if the maintenance resource is at Asset 1, it is optimal to perform proximal maintenance.

4.4.2 Positioning and Deterioration Conditions

To understand how the deterioration vector $\boldsymbol{x}$ affects the value function and the corresponding optimal actions at different locations, we look at two scenarios: (i) when deterioration levels are different among assets, i.e., are unbalanced, and (ii) when deterioration levels are equal among all assets, i.e., are balanced.

Unbalanced Deterioration Levels. When deterioration conditions differ among the assets, it is often optimal to move toward the assets in higher levels of deterioration. However, such policies are not necessarily optimal in general graph structures, especially when assets are not located at equally central nodes. For instance, it may be optimal to idle or move toward less deteriorated assets if the maintenance resource is close to those locations. That is, the maintenance resource would take advantage of its proximity to these assets to perform early maintenance or idle at these locations in anticipation of further changes in their conditions. Figure 16 depicts an example of such a scenario for $c_T = 1$, $c_R(0) = 0.2$, $c_R(1) = 0.4$, $c_R(2) = 0.6$, $c_R(3) = 1$, $c_D = 1$ and

$$P = \begin{bmatrix} 0.8 & 0.15 & 0 & 0.05 \\ 0 & 0.8 & 0.15 & 0.05 \\ 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

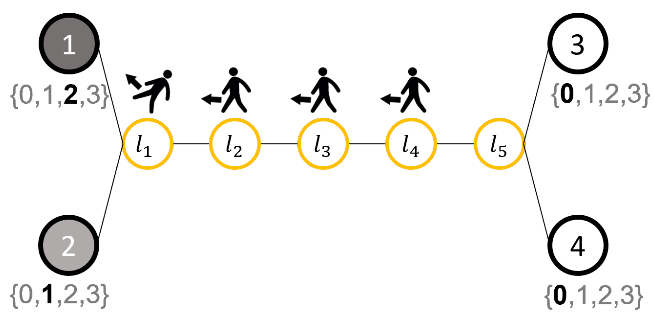


Figure 16: An excerpt of the optimal policy for unbalanced deterioration levels and high travel cost. It is optimal to idle at the locations close to Assets 3 and 4 even though Assets 1 and 2 are more deteriorated.

Balanced Deterioration Levels. When deterioration levels among the assets are balanced, our numerical results suggest that under sufficiently low travel costs and healthy deterioration conditions (collectively among all assets), it is optimal to travel toward the central locations of the graph and idle in anticipation of further changes. Conversely, under sufficiently high travel costs and deterioration conditions, it is optimal to travel toward and idle in the asset locations. Figure 17 illustrates this claim for a (2,4)-banana tree with four assets on the leaf nodes. Specifically, Figure 17 plots the value function and the corresponding optimal actions against each node of the graph under different (balanced) deterioration conditions and travel costs. The graph configuration and the parameter settings in Figure 17 are the same as those in Figure 16, except for the travel costs.

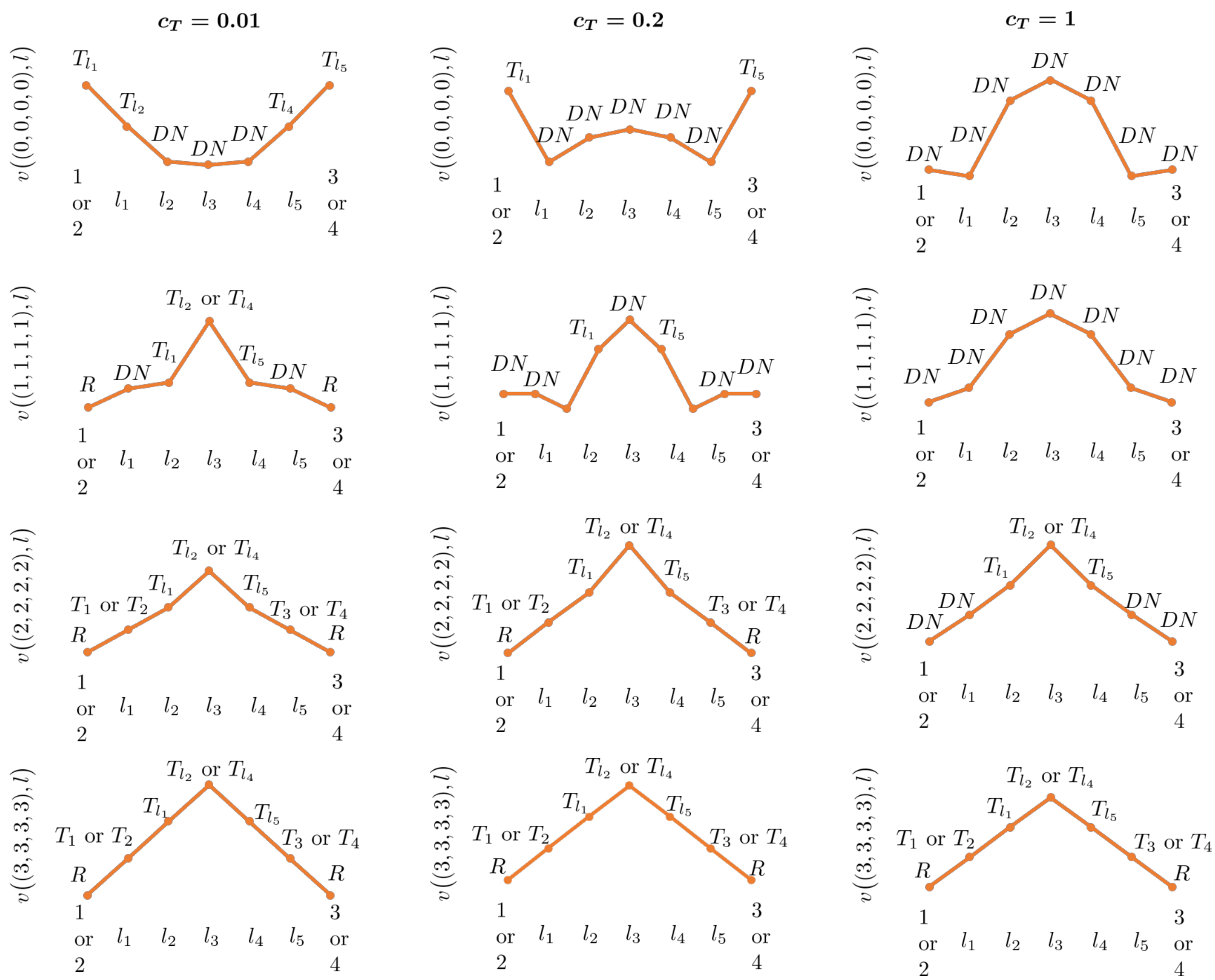


Figure 17: Each plot depicts the value function and the corresponding optimal actions for different locations in a (2,4)-banana tree graph with assets on all four leaf nodes. Travel cost increases left to right and (balanced) deterioration conditions from top to bottom. Under low travel costs and deterioration conditions, it is optimal to move toward and idle in the middle locations of the graph. Conversely, under high travel costs and deterioration conditions, it is optimal to move closer to and idle at the asset locations. Lower travel costs also prompt earlier proximal maintenance. Note that the plots are consistent with the results established in Theorems 10 and 11.

Additionally, note that in all plots of Figure 17, the locations with locally minimum long-run expected discounted cost correspond to idling or repairing actions as established in Theorem 11. Also, for every optimal repositioning action, the long-run expected discounted cost of the corresponding location is larger than that of the destination location (see Theorem 11).

4.4.3 Idling and Graph Centrality

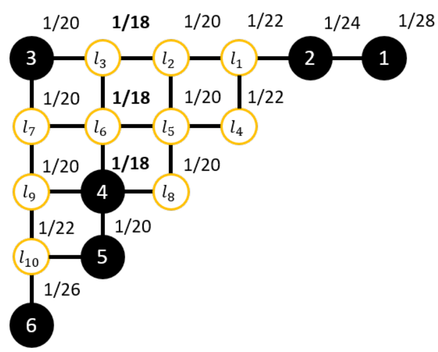
In this section our goal is to identify the locations in the graph that are most used for idling under the optimal policy and employ *graph centrality measures* [98] to explore the connections between these idling locations and graph structure. We simulate the optimal actions of the maintenance resource and its movement through the graph and record the number of time units the maintenance resource (i) spends in each node, or (ii) idles, i.e., implements the do nothing action, in each node. We then report the long-run average fraction of time spent (or idle time spent) at each node and visualize these averages as *heat maps*.

Recall from Section 4.2 that, under our modeling assumptions, one unit of time elapses if the optimal action is to repair an asset, idle (do nothing), or reposition to an adjacent node. For metric (i), we record the cumulative time spent repairing, idling, and traveling through each node. For metric (ii), we only record the time spent idling, which we later exploit in Section 4.6. The heat map for metric (i) can also indicate the most frequently traversed paths on the graph. Examples of heat maps for both metrics are depicted in Figures 18b and 18c, respectively.

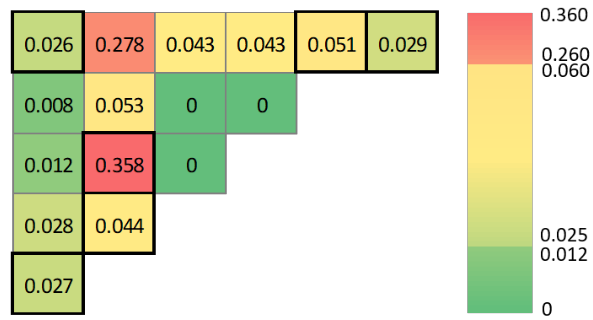
Example 16. Assume six assets dispersed on a graph as depicted in Figure 18a. Parameter values are $\Delta = 2$, $c_R(0) = 1$, $c_R(1) = 2$, $c_R(2) = 3$, $c_D = 2$, $c_T = 0.05$, $\lambda = 0.995$, and

$$P = \begin{bmatrix} 0.98 & 0.01 & 0.01 \\ 0 & 0.98 & 0.02 \\ 0 & 0 & 1 \end{bmatrix}.$$

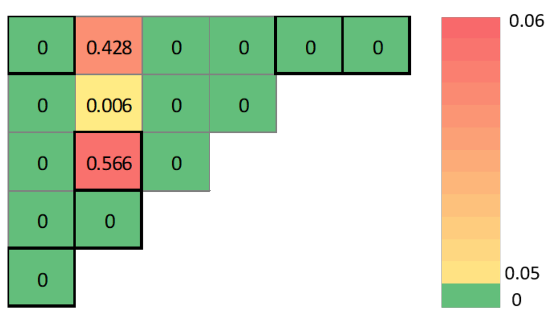
To obtain the average fraction of time spent in each node, we conduct a simulation study to trace the optimal movements of the maintenance resource for 1100000 units of time after a warm up period of 4000 units of time. (Note that here we choose a sufficiently long simulation duration such that the differences in averages obtained in three consecutive runs is less than 0.1%.) In the warm up period, we do not trace the movements so that we can exclude the transient behavior. These averages are then visualized as a heat map in Figure 18b. Similarly, the averages for time spent idling only are visualized in Figure 18c.



(a) Graph configuration of Example 16; assets are located at black nodes. Closeness centrality scores, see equation (4.8), are above and to the right of each node.



(b) Heat map for fraction of overall time spent in each location, and its color scale. Asset locations are depicted with thick borders.



(c) Heat map for fraction of idle time spent in each location.

Figure 18: Idling is only optimal at nodes 4, l_3, and l_6 which are the most central nodes with respect to the closeness centrality measure.

The heat map in Figure 18b illustrates the nodes at which the maintenance resource spends most of its time. Moreover, a comparison of Figures 18b and 18c identifies nodes used only for traveling (i.e., those with zero value in Figure 18c) and thus the regions most frequently traversed under the optimal policy. Notice that the maintenance resource never visits nodes l_4, l_5, and l_8 because alternative paths of the same length exist and are closer to all assets; such nodes can be eliminated.

The heat map in Figure 18c indicates that idling is optimal only in three nodes: Asset 4, l_3, and l_6. This observation holds across a wide range of parameter values (not presented here). Under the optimal policy, when all assets are as-good-as-new, the maintenance resource travels toward the closest idling node (i.e., 4, l_3, or l_6), or idles if it is already at one of these nodes. When an asset slightly deteriorates (i.e., reaches deterioration level 1), the maintenance resource travels toward the idling node that is closest to that asset (or idles if it is located at that node). Once an asset is sufficiently deteriorated (i.e., reaches deterioration level 2), the maintenance resource travels toward and maintains that asset.

Our numerical work and the examples in this section suggest that idling nodes are affected by the graph structure and tend to be centrally positioned with respect to the asset nodes. In graph theory and network analysis, centrality is a fundamental concept to identify the most "important" nodes within a graph. Various measures have been proposed that use different definitions of centrality to identify such important nodes; examples include degree, Eigenvector, Katz, closeness, and betweenness centrality. These measures reflect different aspects of connectivity and are real-valued functions that provide a ranking of each node with respect to the centrality measure. For instance, degree centrality is characterized by the number of links incident upon a node; Eigenvector and Katz centralities measure the influence of a node based on connections to high-scoring nodes; closeness centrality measures how close a node is on average to other nodes; and, betweenness centrality quantifies how often a node acts as a bridge on the shortest path between other nodes [98]. Next, we propose two measures inspired by closeness and betweenness centralities that we believe are well-suited measures to identify idling nodes.

We define a closeness centrality measure as

$$\mathcal{C}(l) = \frac{1}{\sum\limits_{i \in V_M} d_G(i,l)}, \tag{4.8}$$

where $d_G(i,l)$ denotes the length of the shortest path between nodes i and l. Note that in the denominator of (4.8) we only include *asset* nodes in the summation of graph distances because, in our application, the maintenance resource is positioned and travels between assets. In a typical network science application, however, the denominator may include all nodes [98]. Equation (4.8) assigns larger scores to nodes that are closer to all asset locations.

Interestingly, in Example 16, idling nodes 4, l_3, and l_6 have the largest closeness centrality score (i.e., 1/18) among all nodes; see Figure 18a. This example demonstrates the relationship between the optimal policy obtained through the MDP formulation and the graph structure that connects asset nodes, and that this relationship can be explained by appropriate graph centrality measures such as closeness centrality. ■

Our numerical work also suggests that betweenness centrality together with closeness centrality can identify idling locations. Example 17 illustrates the relationship between idling locations and their closeness and betweenness centrality measures.

Example 17. Assume four assets dispersed on a graph as shown in Figure 19a. We let $\Delta = 3$, $c_R(0) = 2$, $c_R(1) = 2$, $c_R(2) = 2$, $c_R(3) = 3$, $c_D = 2$, $c_T = 0.1$, $\lambda = 0.995$, and

$$P = \begin{bmatrix} 0.98 & 0.01 & 0.01 & 0 \\ 0 & 0.98 & 0.01 & 0.01 \\ 0 & 0 & 0.98 & 0.02 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

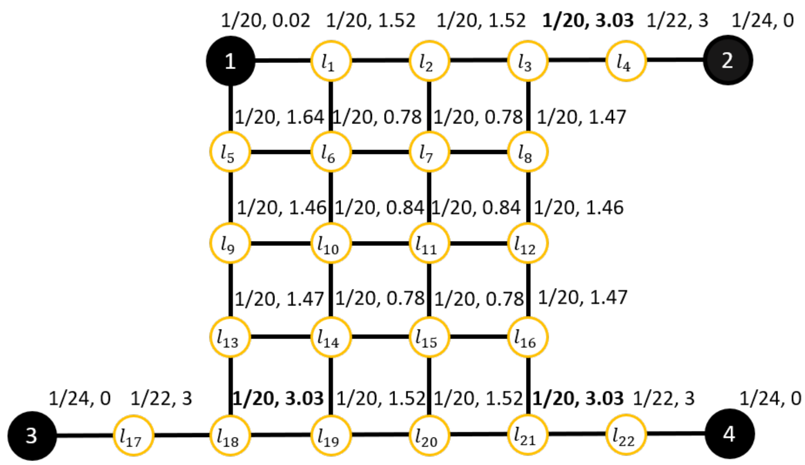
We run a simulation study to trace the optimal movements of the maintenance resource for 1400000 units of time after a warm up period of 40000 units of time. The resulting heat maps are presented in Figures 19b and 19c.

In our application, we define betweenness centrality as

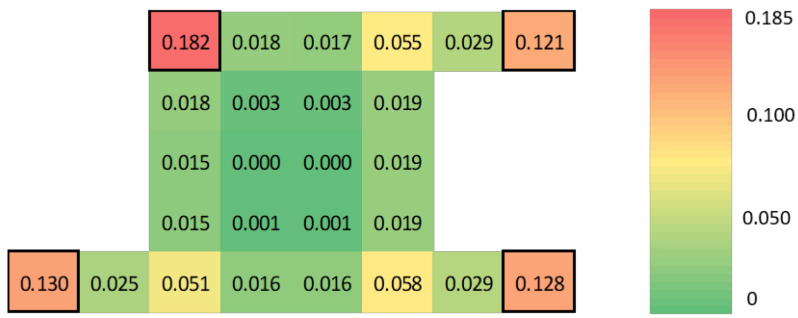
$$\mathcal{B}(l) = \sum_{i,j \in V_M, i \neq j \neq l} \frac{\sigma_{ij}(l)}{\sigma_{ij}}, \tag{4.9}$$

where σ_{ij} is the total number of shortest paths from node i to node j and $\sigma_{ij}(l)$ is the number of those paths that pass through l. Similar to the closeness centrality measure, we only include paths between assets. Equation (4.9) assigns larger scores to nodes that are located on a larger number of shortest paths between assets.

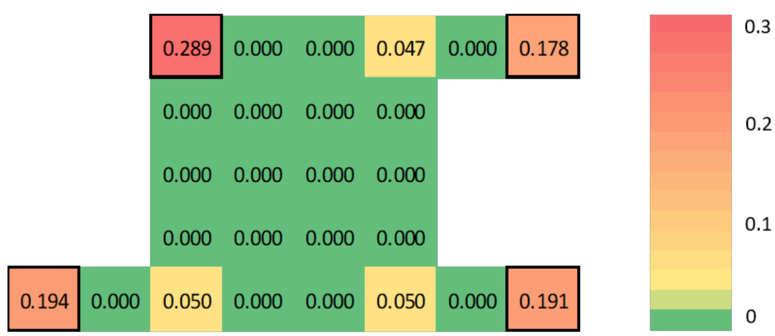
In Example 17, the middle twenty nodes have equal closeness centrality scores. Among these central locations, nodes 1, l_3, l_{18}, and l_{21} have the largest betweenness centrality scores; that is, most of the shortest paths between assets pass through these nodes. When all assets are as-good-as-new, the maintenance resource moves toward and idles at the closest one of these nodes. Otherwise, it moves toward a deteriorated asset. ■



(a) Graph configuration of Example 17; assets are located at black nodes. For each node, closeness centrality and between centrality scores, see equations (4.8) and (4.9), are above and to the right, respectively.



(b) Heat map for fraction of overall time spent in each location, and its color scale. Asset locations are depicted with thick borders.



(c) Heat map for fraction of idle time spent in each location.

Figure 19: Idling is optimal at asset nodes and at auxiliary nodes l_3, l_{18}, and l_{21}. Among the nodes with the largest closeness centrality score, these auxiliary nodes have the largest betweenness centrality scores.

In summary, Examples 16 and 17 highlight the connection between our problem setting and graph structure. In general, when assets are not sufficiently deteriorated, it is optimal to move toward central locations and idle there in anticipation of further deterioration. Another takeaway from these examples is that our simulations and the resulting heat maps can provide a holistic view of the maintenance resource's optimal movements and identify the important nodes and paths within a graph. Maintenance planners can potentially use such "important" locations as bases, resource and material storage depots, and battery charging stations in robotic applications.

Moreover, we find that graph theory measures such as closeness and betweenness centrality can be used to identify candidate locations *a priori*. The close relationship between these measures and the idling locations obtained via simulating the optimal policy is intuitive in that they identify nodes that are close to all assets and bridge most paths traversed between them. In Section 4.6 we use these measures to generate easy-to-implement heuristic policies.

4.5 Additional Performance Metrics

Before exploring heuristic policies, in this section we conduct a simulation study to quantify several metrics of interest under different parameter values to derive managerial insights for maintenance planners. More specifically, we consider the average percentage of time that the maintenance resource spends performing repairs (proximal and/or reactive; i.e., at failure), traveling, or idling (in auxiliary and/or asset locations); and, the average percentage of time that the assets are down.

Consider two network configurations, namely, a (2,4)-banana tree with four assets on the leaf nodes, and a (3,3)-grid graph with four assets on the corner nodes (Tables 10 and 11, respectively). Note that in both graphs, the average percentage of downtime among all assets is nearly equivalent in the long-run because they are in equally central locations (i.e., they have equal closeness centrality measures). Hence, we only report the average percentage of downtime across the four assets. See Appendix C for additional graph configurations and a comparison of downtime among assets with different closeness centrality measures.

To assess how the parameter values affect the metrics of interest, first, Tables 10 and 11 report the values of the metrics for a *benchmark* scenario (second column). Under the benchmark scenario, we let $\Delta = 3$, $c_R(0) = 0.1$, $c_R(1) = 0.2$, $c_R(2) = 0.3$, $c_R(3) = 0.5$, $c_D = 0.1$, $c_T = 0.01$, $\lambda = 0.995$, and

$$P = \begin{bmatrix} 0.95 & 0.03 & 0.01 & 0.01 \\ 0 & 0.95 & 0.03 & 0.02 \\ 0 & 0 & 0.95 & 0.05 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

In the remaining columns of Tables 10 and 11, we change the value of a single parameter, while keeping all other parameter values equal to the benchmark scenario. Thus, in each column we can evaluate the effect of changing the corresponding parameter value on our metrics of interest.

Regarding the transition probabilities, we let P vary as follows:

$$P' = \begin{bmatrix} 0.9 & 0.06 & 0.03 & 0.01 \\ 0 & 0.9 & 0.06 & 0.04 \\ 0 & 0 & 0.9 & 0.1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; P'' = \begin{bmatrix} 0.85 & 0.1 & 0.03 & 0.02 \\ 0 & 0.85 & 0.1 & 0.05 \\ 0 & 0 & 0.85 & 0.15 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \text{ and}$$

$$P''' = \begin{bmatrix} 0.8 & 0.15 & 0 & 0.05 \\ 0 & 0.8 & 0.15 & 0.05 \\ 0 & 0 & 0.8 & 0.2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

These transition probability matrices (P, P', P'', P''') have the IFR property and are upper triangular; note that we impose upper triangularity because in many applications it is unlikely that the deterioration conditions of the assets improve in the absence of maintenance interventions. Moreover, to explore the effect of deterioration speed on the optimal policy, we construct the transition matrices such that they are orderable under first-order stochastic dominance [78, 121], i.e., $P \preceq P' \preceq P'' \preceq P'''$. The stochastic dominance of P' over P, for instance, implies that in any given time period, given the current deterioration condition it is more likely for an asset to deteriorate to a worse condition under P' compared to P.

Table 10: The effect of changing parameter values on metrics of interest in a (2,4)-banana tree with assets on all leaf nodes. All metrics are presented in percent values. The benchmark scenario represents the metric values under $\Delta = 3$, $c_R(0) = 0.1$, $c_R(1) = 0.2$, $c_R(2) = 0.3$, $c_R(3) = 0.5$, $c_D = 0.1$, $c_T = 0.01$, and $\lambda = 0.995$. In the columns under "Repair cost," the + sign indicates adding 0.1 to the benchmark values of $c_R(0), c_R(1), c_R(2), c_R(3)$. "Maintenance and Traveling" is equivalent to the percentage of time that the maintenance resource is either travelling or performing maintenance; i.e., $(1 - \text{idle})\%$. "Downtime" is equivalent to the average percentage of time assets are down.

		Repair cost			Downtime cost			Travel cost			Transition matrix			Discount factor		
		$c_R(0)$,...,$c_R(\Delta)$			c_D			c_T			P			λ		
	benchmark	+0.1	+0.5	+1	0.5	2	6	0.05	0.1	0.5	P'	P''	P'''	0.99	0.95	0.8
Maintenance:	10.1	9.8	8.5	8.5	10.7	10.9	10.9	9.6	9.8	7.8	17.6	21.1	22.8	9.9	9.1	8.5
Preventive (before failure)	3.2	2.4	0	0	4.3	4.7	4.8	2.3	2.4	1.3	5.8	4.7	2.8	2.6	1.1	0
Reactive	6.9	7.4	8.5	8.5	6.4	6.2	6.2	7.3	7.3	6.5	11.8	16.4	20.0	7.3	8.0	8.5
Travel	45.1	44.4	42.3	41.4	52.7	56.9	58.4	36.3	31.3	16.2	69.4	75.2	76.0	44.0	42.1	39.9
Idle:	44.8	45.7	49.2	50.2	36.6	32.2	30.7	54.1	60.0	75.9	13.0	3.8	1.2	46.2	48.9	51.6
Auxiliary locations	26.5	23.6	43.4	49.7	32.4	32.2	30.7	21.9	2.5	0.0	8.5	2.4	0.7	24.2	35.4	41.9
Asset locations	18.3	22.1	5.7	0.42	4.2	0	0	32.1	56.4	75.9	4.5	1.4	0.5	22.0	13.5	9.7
Maintenance and Traveling	55.2	54.3	50.8	49.8	63.4	67.8	69.3	45.9	41.0	24.1	87.0	96.2	98.8	53.8	51.1	48.4
Downtime	8.5	8.8	9.9	10.2	8.2	8.4	8.3	9.6	10.5	26.1	18.5	29.8	40.2	8.8	9.4	10.0

Table 11: The effect of changing parameter values on metrics of interest in a (3,3)-grid graph with assets on all corner nodes. All metrics are presented in percent values. The benchmark scenario represents the metric values under $\Delta = 3$, $c_R(0) = 0.1$, $c_R(1) = 0.2$, $c_R(2) = 0.3$, $c_R(3) = 0.5$, $c_D = 0.1$, $c_T = 0.01$, and $\lambda = 0.995$. In the columns under "Repair cost," the + sign indicates adding 0.1 to the benchmark values of $c_R(0), c_R(1), c_R(2), c_R(3)$. "Maintenance and Traveling" is equivalent to the percentage of time that the maintenance resource is either travelling or performing maintenance; i.e., $(1 - \text{idle})\%$. "Downtime" is equivalent to the average percentage of time that assets are down.

		Repair cost			Downtime cost			Travel cost			Transition matrix			Discount factor		
		$c_R(0),\ldots,c_R(\Delta)$			c_D			c_T			P			λ		
	benchmark	+0.1	+0.5	+1	0.5	2	6	0.05	0.1	0.5	P'	P''	P'''	0.99	0.95	0.8
Maintenance:	9.7	8.9	8.8	8.8	10.7	10.8	10.8	9.6	9.6	9.5	19.3	25.3	29.6	9.3	8.8	8.8
Preventive (before failure)	1.8	0.1	0	0	3.9	3.9	3.9	1.6	1.8	2.1	6.6	6.6	5.3	0.9	0	0
Reactive	7.9	8.8	8.8	8.8	6.8	6.9	6.9	8.0	7.8	7.4	12.7	18.7	24.3	8.4	8.8	8.8
Travel	28.7	28.0	28.1	27.9	35.9	36.5	39.7	21.4	17.9	13.1	50.8	56.8	60.4	28.3	26.5	20.5
Idle:	61.6	63.1	63.1	63.2	53.4	52.7	49.5	68.9	72.5	77.4	29.9	17.9	10.0	62.4	64.7	70.8
Auxiliary nodes	0.7	1.0	1.0	1.4	1.0	3.2	16.3	0.7	0.0	0.0	5.6	0.2	0.0	0.9	1.1	0.3
Asset nodes	60.8	62.1	62.1	61.8	52.3	50.0	33.2	68.2	72.5	77.4	24.3	17.7	10.0	61.6	63.6	70.5
Maintenance and Traveling	38.4	36.9	36.9	36.8	40.6	47.3	50.5	31.1	27.5	22.6	70.1	82.1	90.0	37.6	35.3	29.2
Downtime	5.2	6.0	6.2	6.2	4.7	4.7	4.5	5.9	6.5	11.7	9.8	17.1	27.0	5.7	6.3	7.0

Examining the performance metrics in Tables 10 and 11 generates several insights. First, note that because the distances between the asset locations are smaller in the grid graph compared to the banana tree, idling time is lower in the latter. That is, the maintenance resource spends more time performing maintenance and traveling. However, despite more maintenance and traveling, the percentage of downtime is also higher in the banana tree compared to the grid graph.

Second, note that as repair or travel costs increase, the percentage of time spent performing maintenance and traveling decreases, and conversely, the percentage of time idling increases. This effect is reversed under high downtime costs and discount factors, as well as faster deterioration. These intuitive observations remain consistent across all scenarios in both graph configurations and suggest that increases in maintenance are generally associated with increases in traveling. However, changes in the percentage of time spent idling at assets versus auxiliary locations are somewhat arbitrary and are not consistent in the two graph configurations. For instance, in the grid graph, the percentage of time spent idling in auxiliary locations increases in downtime cost, whereas in the banana tree, the change is non-monotone.

Lastly, note that asset downtime remains consistent under different values of the downtime cost seemingly because maintenance actions also cause downtime. On the other hand, asset downtime notably varies under different deterioration behavior and travel costs.

To summarize the discussion above:

(i) The maintenance resource spends more time idling in better connected graphs with smaller distances between assets.

(ii) Assets incur less downtime in better connected graphs with smaller distances between them. Hence, maintenance planners may save costs by utilizing multiple maintenance resources, each responsible for the maintenance activities of a well-connected cluster of assets.

(iii) Percentage of time idling increases in repair or travel costs and conversely decreases in downtime cost, the discount factor, and deterioration speed. Also, increases in maintenance are generally associated with increases in traveling.

4.6 Heuristic Policies

Next, we propose heuristic policies that are appealing from an implementation perspective in that they consist of simple rules of thumb to be implemented by the maintenance resource. Often in practice, once an asset is identified for maintenance, the resource must commit to traveling to and maintaining that asset first, before taking on other tasks. We consider heuristic policies that follow such practices and compare their performance with that of policies that adopt more dynamic routing rules. Following our discussion in Section 4.3, all of the heuristics adhere to the control-limit rule established in Theorem 10.

The remainder of this section is organized as follows. We first formally define the proposed heuristics in Section 4.6.1. We then compare their performance against the optimal policy for small problem instances in Section 4.6.2, and against each other for large problem instances in Section 4.6.3. Note that the state of the MDP grows exponentially in the number of assets, and thus, for large instances we cannot solve our problem to optimality in a practical amount of time using the value iteration algorithm.

4.6.1 Heuristic Policy Definitions

We consider three types of heuristic policies, namely, (*i*) Committed Maintenance, (*ii*) Committed Routing with Proximal Maintenance, and (*iii*) Dynamic Routing with Proximal Maintenance. All three adopt "anticipatory idling" rules that route the maintenance resource to the central locations of the graph when assets are sufficiently healthy. We refer to these central nodes as "anticipatory idling nodes" and define them by the set A. To obtain A, we choose the n_A nodes with the largest betweenness centrality score among the $2n_A$ nodes with the largest closeness centrality score (recall our discussion in Section 4.4.3). Later in Section 4.6.3 we discuss how we select the value of n_A and explore alternative approaches for obtaining A.

Committed Maintenance. Define the *travel threshold* τ^* such that once the condition of an asset reaches (or exceeds) τ^*, the heuristic action is to travel to the location of that asset and maintain it upon arrival. In case of ties, the maintenance resource chooses the closest asset. This rule implies that the maintenance resource cannot change course en route to the destination asset and commits to performing maintenance upon arrival. This approach captures common practices where sufficiently deteriorated assets are identified for maintenance and work orders are made for, e.g., human crews to travel to and repair those assets. We also define an *idling threshold* ι^* such that if all assets are healthier than that threshold, then the heuristic action is to travel toward the closest *anticipatory idling node* within the set of nodes in A. We formalize our policy definition as follows:

$$a_{CM}(\boldsymbol{x}, l) = \begin{cases} R & \text{if } l \text{ is a destination asset and } x_l \geq \tau^*, \\ DN & \text{if } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^* \text{ and } l \in A, \\ T_b & \text{if } l \text{ is not a destination asset, } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^*, \\ & \quad \text{and } (l, b) \text{ is on the shortest path to} \\ & \quad \text{(the closest) anticipatory idling node in } A, \\ T_b & \text{if } l \text{ is not a destination asset, } \max\limits_{j \in \{1,...,n_M\}} x_j > \iota^*, \\ & \quad \text{and } (l, b) \text{ is on the shortest path to (the closest) destination asset,} \end{cases} \quad \text{for } l \in V_M,$$

and,

$$a_{CM}(\boldsymbol{x}, l) = \begin{cases} DN & \text{if } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^* \text{ and } l \in A, \\ T_b & \text{if } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^* \text{, and } (l, b) \text{ is} \\ & \quad \text{on the shortest path to (the closest)} \\ & \quad \text{anticipatory idling node in } A, \\ T_b & \text{if } \max\limits_{j \in \{1,...,n_M\}} x_j > \iota^* \text{, and } (l, b) \text{ is} \\ & \quad \text{on the shortest path to (the closest) destination asset,} \end{cases} \quad \text{for } l \in V_T,$$

where $a_{CM}(\boldsymbol{x}, l)$ denotes the action prescribed under the Committed Maintenance heuristic.

To determine the values of τ^* and ι^*, we compute the long-run average cost-per-unit time incurred under all combinations of their values via simulation and choose a combination that yields the lowest cost-per-unit time. We let $\tau \in \{1, 2, ..., \Delta\}$ and $\iota \in \{-1, 0, 1, ..., \tau - 1\}$, where $\iota = -1$ indicates a policy with no anticipatory idling rules; i.e., the maintenance resource only travels between and idles at asset nodes. Our numerical experiments indicate that the cost-per-unit time is unimodal with respect to the value of ι^* for a fixed τ^*, indicating that we can stop the search once we reach a local minimum; we did not observe any instances where this unimodality was violated.

Committed Routing with Proximal Maintenance. This policy is similar to the previous one in that the maintenance resource uses a *travel threshold* τ^* to identify and travel to an asset, and a *idling threshold* ι^* to reposition and idle at central locations. However, once the resource arrives at the destination asset (or an asset en route to the destination asset), the maintenance resource uses a *maintenance threshold* μ^* to make a repair decision. That is, at the destination asset, the heuristic action is to repair if the deterioration condition of the asset is μ^* or greater and is to idle otherwise. This rule also implies that the heuristic action is to repair if the maintenance resource arrives at an asset with a deterioration condition μ^* or greater on its way to the destination asset; we refer to this scenario as proximal maintenance (recall Section 4.4.1). We have that $\mu^* \geq \tau^* > \iota^*$.

To determine the threshold values, we first let $\iota = -1$ and compute the long-run average cost-per-unit time incurred under all combinations of μ and τ via simulation and choose a combination that yields the lowest cost-per-unit time. We let $\tau \in \{1, ..., \mu\}$ and $\mu \in \{\tau, ..., \Delta\}$. We then determine ι^* by enumerating the values of $\iota \in \{-1, 0, ..., \tau^* - 1\}$.

We formalize our definition below:

$$a_{CR}(\boldsymbol{x}, l) = \begin{cases} R & \text{if } x_l \geq \mu^*, \\ DN & \text{if } l \text{ is a destination asset and } x_l < \mu^*, \\ DN & \text{if } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^* \text{ and } l \text{ is} \\ & \text{an anticipatory idling node in } A, \\ T_b & \text{if } l \text{ is not a destination asset, } \max\limits_{j \in \{1,...,n_M\}} x_j \leq \iota^*, \\ & \text{and } (l,b) \text{ is on the shortest path to (the closest)} \\ & \text{anticipatory idling node in } A, \\ T_b & \text{if } l \text{ is not a destination asset, } \max\limits_{j \in \{1,...,n_M\}} x_j > \iota^*, \text{ and } (l,b) \\ & \text{is on the shortest path to (the closest) destination asset,} \end{cases} \quad \text{for } l \in V_M,$$

and,

$$a_{CR}(\boldsymbol{x}, l) = \begin{cases} DN & \text{if } \max\limits_{j\in\{1,\dots,n_M\}} x_j \leq \iota^* \text{ and } l \in A, \\ T_b & \text{if } \max\limits_{j\in\{1,\dots,n_M\}} x_j \leq \iota^* \text{ , and } (l,b) \text{ is} \\ & \quad \text{on the shortest path to (the closest)} \\ & \quad \text{anticipatory idling node in } A, \\ T_b & \text{if } \max\limits_{j\in\{1,\dots,n_M\}} x_j > \iota^* \text{ , and } (l,b) \text{ is} \\ & \quad \text{on the shortest path to (the closest) destination asset,} \end{cases} \quad \text{for } l \in V_T,$$

where $a_{CR}(\boldsymbol{x}, l)$ denotes the action prescribed under the Committed Routing with Proximal Maintenance heuristic.

Dynamic Routing with Proximal Maintenance. Here, the maintenance and anticipatory idling rules are the same as those described under the Committed Routing with Proximal Maintenance policy. However, destination nodes are updated at each decision epoch (i.e., upon arrival at each node or completing a repair action) such that the resource is routed to the asset with the worst deterioration condition. Thus, the maintenance resource may change course en route to a previously selected destination node. The rules are formally defined below:

$$a_{DR}(\boldsymbol{x}, l) = \begin{cases} R & \text{if } x_l \geq \mu^*, \\ DN & \text{if } x_l = \max\limits_{j\in\{1,\dots,n_M\}} x_j > 0 \text{ and } x_l < \mu^*, \\ T_b & \text{if } \max\limits_{j\in\{1,\dots,n_M\}} x_j > x_l \text{ and edge } (l,b) \text{ is} \\ & \quad \text{on the shortest path to (the closest) } \operatorname*{arg\,max}\limits_{j\in\{1,\dots,n_M\}} x_j, \\ T_b & \text{if } \max\limits_{j\in\{1,\dots,n_M\}} x_j \leq \iota^* \text{ and } (l,b) \text{ is} \\ & \quad \text{on the shortest path to (the closest) anticipatory idling node in } A, \end{cases} \quad \text{for } l \in V_M,$$

and,

$$a_{DR}(\boldsymbol{x}, l) = \begin{cases} DN & \text{if } \max\limits_{j \in \{1,\dots,n_M\}} x_j \leq \iota^* \text{ and } l \in A, \\ T_b & \text{if } \max\limits_{j \in \{1,\dots,n_M\}} x_j > \iota^* \text{ and edge } (l,b) \text{ is} \\ & \quad \text{on the shortest path to (the closest) } \operatorname*{arg\,max}\limits_{j \in \{1,\dots,n_M\}} x_j, \\ T_b & \text{if } \max\limits_{j \in \{1,\dots,n_M\}} x_j \leq \iota^*,\ l \notin A, \\ & \quad \text{and edge } (l,b) \text{ is on the shortest path to} \\ & \quad \text{(the closest) anticipatory idling node in } A, \end{cases} \qquad \text{for } l \in V_T,$$

where $a_{DR}(\boldsymbol{x}, l)$ denotes the action under the Dynamic Routing with Proximal Maintenance heuristic. Similar to the previous heuristics, the values of μ^* and ι^* are again optimized via enumeration.

Note. In each of the heuristic policies, the maintenance resource traverses the shortest path between the current node and the "destination node" (e.g., the most deteriorated asset or an auxiliary idling node). If there are multiple shortest paths, then one of them is chosen arbitrarily.

4.6.2 Optimality Gap

Here, we conduct a numerical study to assess the performance of the heuristic policies against the optimal policy. We use average cost-per-unit time as a comparison metric, which we compute using simulation. Note that because our Markov decision model is multichain [106], we cannot formulate an average cost-per-unit time optimization model, and instead, derive our optimal policies from the total expected discounted cost-to-go criterion in (4.4). When the discount factor is sufficiently large, the optimal policies under either criterion should coincide [106].

In this study, first, we randomly generate ten instances of four asset locations on a (5,5)-grid graph and one anticipatory idling location; i.e., $n_A = 1$. Then, we compute the percent increase in the cost-per-unit time incurred under each heuristic policy with respect to that under the optimal policy for each random instance. Table 12 reports the mean of the percent increases across the ten random instances under the following parameter settings.

In particular, the second column of Table 12 reports these means for a benchmark scenario under which $\Delta = 3$, $c_R(i) = 40$ for all $i \in \mathcal{K}$, $c_D = 80$, $c_T = 0.1$. In the remaining columns of Table 12, we change the value of a single parameter while keeping all other parameter values equal to the benchmark scenario. The transition probabilities are equal to those in the numerical study of Section 4.5 where we have $P \preceq P' \preceq P'' \preceq P'''$. In all scenarios we assume $\lambda = 0.995$. We run our simulations for forty thousand time units and record the costs after a warm-up period of five thousand time units.

Table 12 indicates that under all settings the Dynamic Routing policy significantly outperforms the other policies that adopt less dynamic routing rules. Moreover, allowing proximal maintenance and relying on a repair threshold (i.e., μ^*) significantly improves policy performance; that is, under all settings, the Committed Routing policy outperforms the Committed Maintenance policy. The performance gap between the heuristics only decreases when deterioration speed increases. This intuitive observation holds because when assets deteriorate rapidly (i.e., under P''') the rules under the heuristics become similar in that the maintenance resource travels often between assets and repairs them when they are highly deteriorated.

Table 12: Percent increase in the mean cost-per-unit time of the proposed heuristic policies against the optimal policy. The percentages are averages across ten random graph instances described in Section 4.6.2. "CM," "CR," and "DR" denote the heuristic policies as Committed Maintenance, Committed Routing and Dynamic Routing with Proximal, respectively.

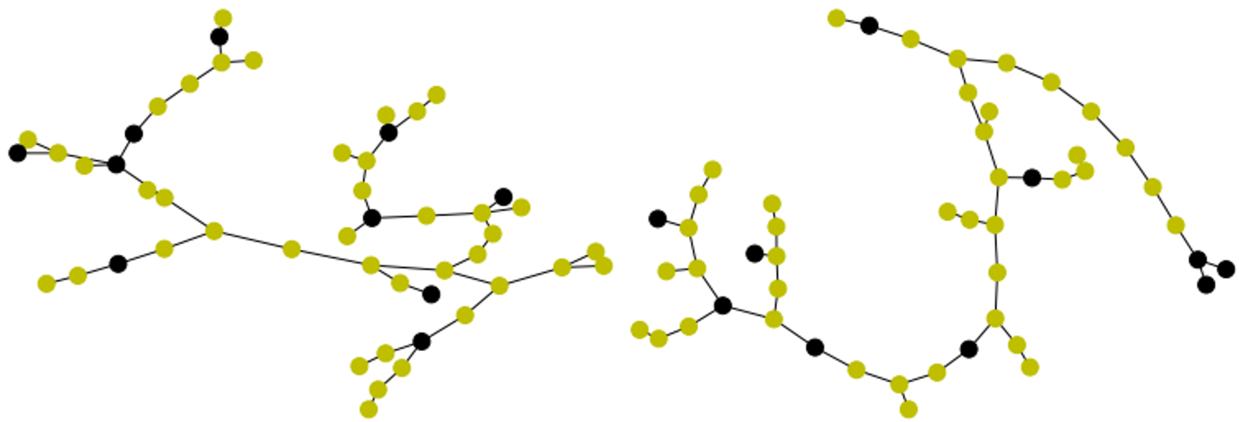
		repair cost			downtime cost			travel cost			transition		
	benchmark	$c_R(0)$,...,$c_R(\Delta)$			c_D			c_T			matrix P		
	scenario	60	100	160	100	120	160	1	5	10	P'	P''	P'''
CM	521%	980 %	518%	252%	725%	889%	735%	720%	741%	498%	321%	227%	85%
CR	276%	270 %	270%	238%	267%	228%	421%	241%	472%	274%	74%	27%	23%
DR	8%	8%	6%	3%	7%	5%	7%	7%	4%	6%	5%	4%	4%

4.6.3 Large Problem Instances

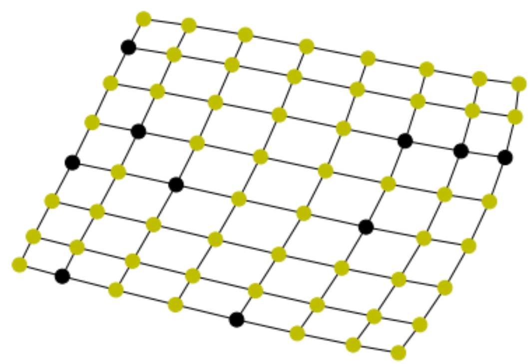
Next, we conduct a numerical study to assess the performance of the heuristic policies against each other for large problem instances where we cannot solve our problem to optimality within a matter of hours. We consider two network configurations, namely, a tree graph with fifty nodes and an (8,8)-grid graph; each instance has ten assets located on randomly selected nodes. We generate the tree graphs using random graph generator functions built in NetworkX, a Python graph library. Figure 20 depicts instances of two tree graphs and a grid graph used in our study.

Similar to the previous section, we randomly generate ten instances of each graph configuration. Then, we compute the cost-per-unit time incurred by implementing each heuristic policy. Figures 21 and 22 report the mean cost-per-unit time across the ten random instances. In particular, in each chart we change the value of a single parameter while keeping all other parameter values equal to a benchmark scenario and depict the mean cost-per-unit time incurred under each policy. Our parameter values are the same as in Section 4.6.2 with the exception that here we let $\Delta = 9$ and change the deterioration probability matrices accordingly. These matrices are presented in Appendix D and preserve the first-order stochastic dominance relationship, i.e., $P \preceq P' \preceq P'' \preceq P'''$. Moreover, we find that the heuristics perform generally well when we assume four anticipatory idling nodes; that is, we let $n_A = 4$. We run our simulations for two hundred thousand time units and record the costs after a warm-up period of one thousand time units.

The conclusions derived from Figures 21 and 22 are consistent with our findings in Table 12 in that the Dynamic Routing policy incurs the smallest average cost-rate, followed by the Committed Routing and Committed Maintenance policies. Also, the performance gap between these policies decreases as deterioration speed increases (see Figures 21d and 22d). Furthermore, note that the average cost-rate under the Committed and Dynamic Routing policies are generally increasing in each cost value. However, the average cost-rate under the Committed Maintenance policy does decreasing in certain cost values in some settings; e.g., see Figures 21c, 22b, and 22c). For large instances with many assets, the Committed Maintenance policy often fails to balance the trade-offs between maintaining



(a) Two randomly generated tree graphs



(b) An (8,8)-grid graph

Figure 20: Instances of graphs used in the numerical study of Section 4.6.3. Yellow nodes indicate auxiliary locations and black nodes indicate asset locations. The tree configuration and asset locations are generated randomly.

late, which may result in the failure of other assets, and maintaining early and too often. Therefore, compared to the other policies, the performance of the Committed Maintenance heuristic is more sensitive to the problem instance.

In the remainder of this section, first we propose alternative approaches for choosing the set of anticipatory idling nodes A and report the performance of the heuristics under these approaches. Next, we propose a simple modification to the fixed repair threshold μ^* and quantify its cost-saving benefits for the instances in Figures 21 and 22. Recall Section 4.4

where we demonstrate the connections between the optimal policy and the graph structure, in particular the graph centrality measures in (4.8) and (4.9). The following sections build on those findings and highlight the benefits of considering the graph structure in designing heuristic policies.

4.6.3.1 Other Approaches for Obtaining Anticipatory Idling Nodes

Recall that we obtain A (i.e., the set of anticipatory idling nodes), by choosing n_A nodes with the largest betweenness centrality score among the $2n_A$ nodes with the largest closeness centrality score, which is a hybrid approach to incorporating both our proposed centrality measures with an emphasis on the closeness centrality score. This approach is inspired by our findings in Section 4.4.3; in particular, recall equations (4.8) and (4.9) and our discussions following Example 17.

In this section, we consider alternative approaches for obtaining set A and compare the performance of the heuristics under those approaches against the current approach. Figure 23 depicts the mean cost-per-time unit across twenty random tree and grid graph instances under the benchmark parameter values. We consider the alternative approaches as follows. "No idling" indicates $A = \emptyset$; "Closeness" and "Betweenness" indicates selecting n_A nodes with the largest closeness and betweenness centralities, respectively, which we compute using (4.8) and (4.9). Similarly, "Degree," "Eigen," and "Katz," indicates selecting n_A nodes with the largest degree, Eigen, and Katz centrality scores, respectively (recall the brief introduction of these scores in Section 4.4.3). Moreover, "Hybrid-CL" indicates the current approach, and "Hybrid-BW" indicates choosing n_A nodes with the largest closeness centrality score among the $2n_A$ nodes with the largest betweenness centrality score. That is, "Hybrid-BW" prioritizes betweenness centrality contrary to the current approach. Lastly, "Random" indicates selecting n_A nodes randomly.

Figures 23a and 23c suggest that the Committed Maintenance and Dynamic Routing policies perform better by obtaining anticipatory idling nodes using the current approach (i.e., "Hybrid-CL") compared to alternative approaches. Specifically, in the tree graphs, our hybrid approach decreases the mean cost-rate incurred under the Committed Maintenance policy by up to 20% compared to the other centrality measures and by 30% compared to the "no idling" approach. Similarly, our hybrid approach decreases the mean cost-rate incurred under the Dynamic Routing policy by up to 5% compared to the other centrality measures. In the grid graphs, this difference is mitigated seemingly because there are many paths between assets and distances may be smaller in a well-connected grid graph compared to a tree graph. Thus, other centrality measures such as degree centrality can also capture important nodes in a grid graph, whereas closeness and betweenness centralities perform better in a tree graph. We want to point out that our conclusions hold across a wide range of parameter values (as set in Table 12). However, for brevity, we only report the results for the benchmark scenario.

Interestingly, Figure 23b suggests that the Committed Routing policy does not benefit from anticipatory idling rules; i.e., this policy performs better under "no idling" compared to the other approaches. This rather counter-intuitive observation holds because under the Committed Routing policy, the maintenance resource is allowed to idle at the *destination* asset in anticipation of its further deterioration, contrary to the Committed Maintenance policy. Moreover, the maintenance resource cannot leave that asset until completion of the maintenance task, contrary to the Dynamic Routing policy. Hence, this policy performs sufficiently well by only traveling between asset locations and employing those locations as anticipatory idling nodes.

4.6.3.2 Modifying the Repair Threshold

Our results suggest that utilizing a repair threshold for making maintenance decisions upon arriving at an asset node significantly improves the heuristic performance; i.e., Committed and Dynamic Routing policies with proximal maintenance significantly outperform the Committed Maintenance policy across all parameter settings. In this section, we propose a simple modification to the repair threshold and quantify its cost-saving benefits.

Recall our numerical observations in Section 4.4.1 where we conclude that repair thresholds are generally affected by (*i*) the degree of asset centrality, (*ii*) the deterioration conditions of other assets, and (*iii*) the proximity of the maintenance resource that allows for early maintenance (i.e., proximal maintenance). The Committed and Dynamic Routing heuristics introduced in Section 4.6.1 already benefit from allowing proximal maintenance. The modification considered here factors in asset centrality and the condition of other assets. Let $l \in V_M$ be the location of the maintenance resource and $\bar{\mu}_l^*$ be the modified repair threshold for Asset l, where

$$\bar{\mu}_l^* = \max\left\{1, \mu^* - \frac{\max\limits_{j\in\{1,\ldots,n_M\}/l} -\kappa^*}{|\max\limits_{j\in\{1,\ldots,n_M\}/l} -\kappa^*|} - \mathbf{1}_{\{l\in \underline{C}\}} + \mathbf{1}_{\{l\in \overline{C}\}}\right\}; \tag{4.10}$$

the set $\underline{C}$ denotes the quarter of the assets with the smallest closeness centrality scores; conversely, the set $\overline{C}$ denotes the quarter of the assets with the largest closeness centrality scores; and $\kappa^* \in \{1, 2, ..., \Delta\}$ is a threshold value for the condition of the remaining assets, the value of which is optimized via enumeration.

Equation (4.10) modifies the repair threshold as follows. If the asset in the maintenance resource location is relatively (non-)central with respect to the other assets, then its repair threshold is (decreased) increased by one; i.e., the maintenance resource remains idle at the current epoch and repairs that asset later. (Recall from Section 4.4.1 that it is often optimal to repair assets with central locations later than their non-central counterparts.) Moreover, if there are other assets with conditions worse than κ^*, then the repair threshold for Asset l is decreased by one; otherwise it is increased by one. That is, if other assets have elevated conditions, then the maintenance resource repairs Asset l early so that it can consequently travel to other deteriorated assets; otherwise, the proposed modification saves costs by *delaying* maintenance.

Tables 13 and 14 quantify the benefit of adopting the modified repair threshold $\bar{\mu}^*$ for the problem instances generated in Figures 21 and 22, respectively. In particular, these tables report the percent decrease in the mean cost-per-unit time incurred under the Committed and Dynamic Routing policies when adopting the modified repair threshold $\bar{\mu}^*$ against the fixed repair threshold μ^*. Note that our numerical study (not presented here) indicates that closeness centrality is a better measure for obtaining $\overline{C}$ and $\underline{C}$ compared to betweenness centrality or letting $\overline{C} = \underline{C} = \emptyset$.

Tables 13 and 14 suggest that the proposed modification of the repair threshold can significantly improve the performance of the Committed Routing policy. This observation holds because the dynamic repair threshold $\bar{\mu}^*$ takes into account the degree of centrality and the conditions of other assets. The Dynamic Routing policy also benefits, but less significantly than the former policy, seemingly because, even in the absence of a dynamic repair threshold, this policy capitalizes on its dynamic routing rules.

In some applications, maintenance resources may be repositioned by humans or are human technicians themselves. In such applications, the Committed Maintenance and Committed Routing policies may be easier to implement, compared to the Dynamic Routing policy which is more suitable for self-propelled maintenance resources. From our numerical study we conclude that for such applications, the Committed Routing policy in conjunction with proximal maintenance and a dynamic repair threshold performs well across a wide range of parameter settings.

4.7 Concluding Remarks

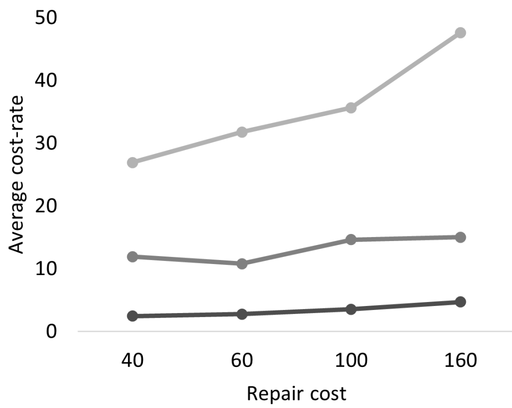
We consider the joint optimization of condition-based maintenance decisions with repositioning and idling decisions of a mobile maintenance resource that is responsible for maintaining a set of geographically distributed assets. We use a graph representation to model possible geographical locations, including idling and asset locations, and the links between them. We formulate a Markov decision process to obtain the optimal travel, idle, and repair actions. To the best of our knowledge, this work is the first to consider dynamic repositioning and condition-based maintenance decisions with idling locations for geographically distributed assets that are maintained by limited number of resources. In the remainder of this section, we provide a summary of our findings and discuss interesting directions for future research.

We establish that under the IFR property and mild conditions on the cost values, the optimal repair action follows a control-limit rule. Moreover, through numerical examples, we illustrate that the optimal repair thresholds are often lower in our geographically dispersed, multi-asset setting compared to the single-asset setting with the same parameter values. This difference result in “early” maintenance in three situations, namely, when an asset’s location is non-central with respect to the other assets; when multiple assets are deteriorated; or when the maintenance resource takes advantage of its proximity to an asset, which we refer to as proximal maintenance.

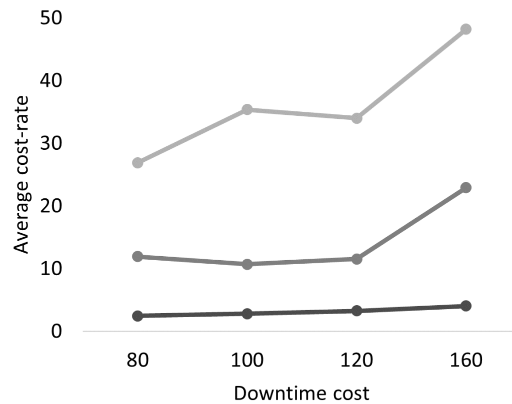
We also observe that when all assets are sufficiently healthy, the maintenance resource often repositions to the central locations of the graph and idles in anticipation of further changes in deterioration conditions. Such central locations can be auxiliary intermediate nodes and are not necessarily asset locations themselves. We show that often these appealing idling locations have close connections to the underlying graph structure. We capture this connection using centrality measures, namely, closeness and betweenness centralities, which are adopted from the network science area. We use these measures to identify critical nodes and generate heuristic policies for large graph settings with many assets where our MDP formulation cannot be solved to optimality (due to the curse of dimensionality).

Moreover, we find that the underlying graph structure that connects asset locations impacts the optimal policy and performance metrics such as resource utilization, asset downtime, and the percentage of time spent performing maintenance. For example, the graph structure affects the maintenance thresholds and the potential anticipatory idling locations. We also find that within a graph, assets in more central locations incur less downtime compared to their non-central counterparts.

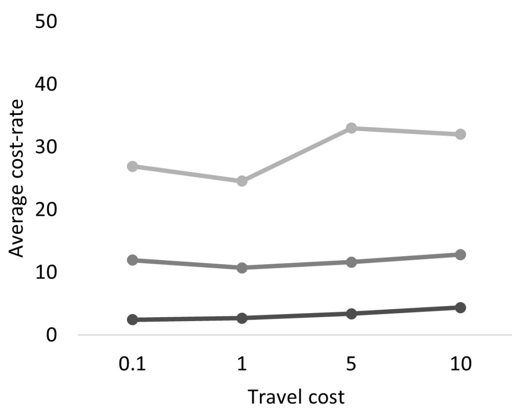
Lastly, we propose easy-to-implement heuristic policies and compare their performances against each other and the optimal policy. We learn that allowing proximal maintenance improves the performance of the policies. Moreover, the heuristic that adopts dynamic routing/positioning rules significantly outperforms those that do not allow repositioning en route to a prespecified asset. Maintenance planners can also adopt repair thresholds that factor in the conditions of the other assets and the centrality score of the asset considered for repair; such modifications especially improve the performance of the Committed Routing policy which may be of interest in applications with human maintenance crews. Lastly, our study indicates that closeness and betweenness centralities, together, yield better candidate anticipatory idling locations compared to alternative centrality scores.



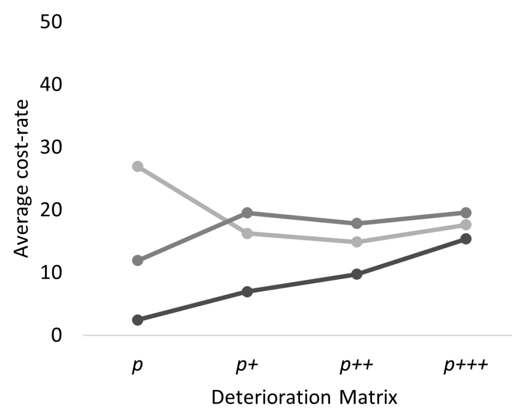
(a) Average cost-rate vs. repair cost



(b) Average cost-rate vs. downtime cost

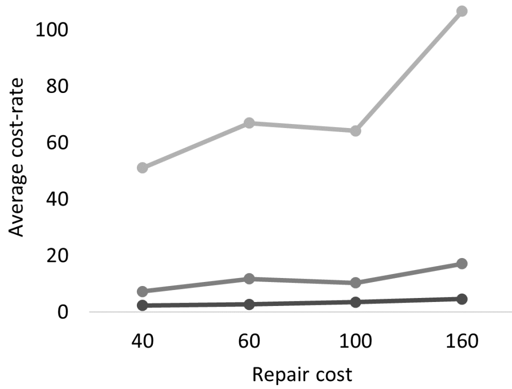


(c) Average cost-rate vs. travel cost

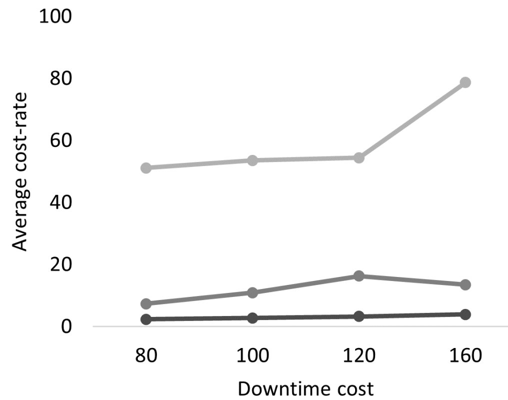


(d) Average cost-rate vs. deterioration matrix

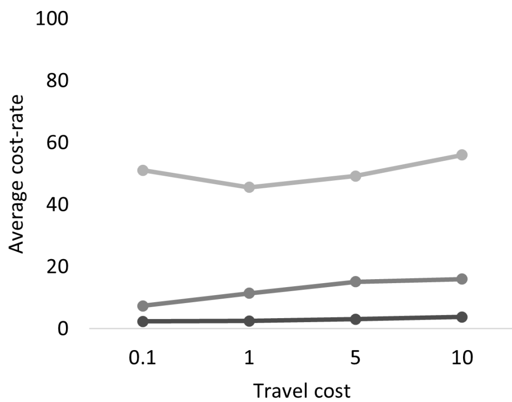
Figure 21: Mean cost-per-unit time under the heuristic policies across ten randomly generated *tree* graphs. In each chart, the value of a single parameter changes while all other parameter values remain equal to the benchmark scenario. Legend colors represent the policies as follows: Committed Maintenance, Committed Routing with Proximal Maintenance, and Dynamic Routing with Proximal Maintenance.



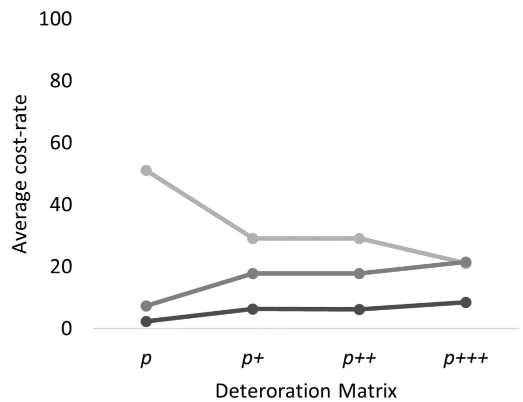
(a) Average cost-rate vs. repair cost



(b) Average cost-rate vs. downtime cost

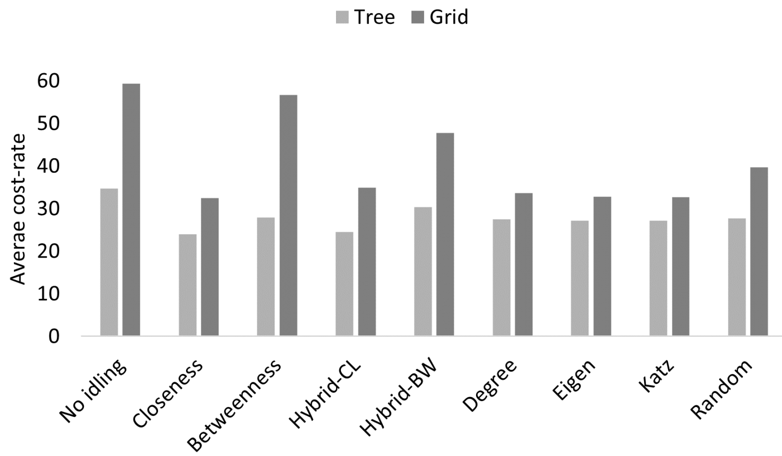


(c) Average cost-rate vs. travel cost

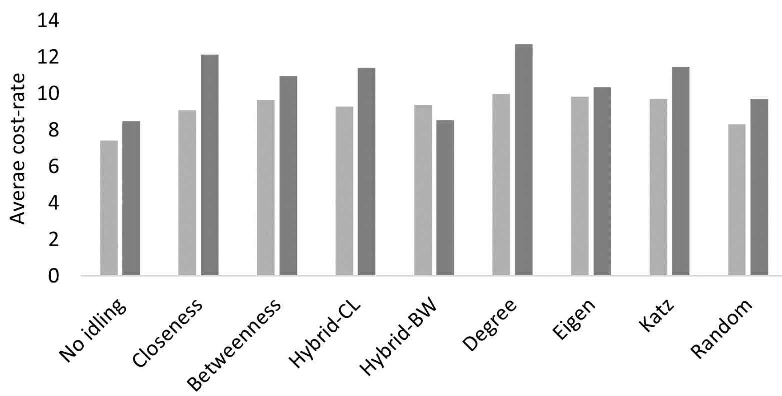


(d) Average cost-rate vs. deterioration matrix

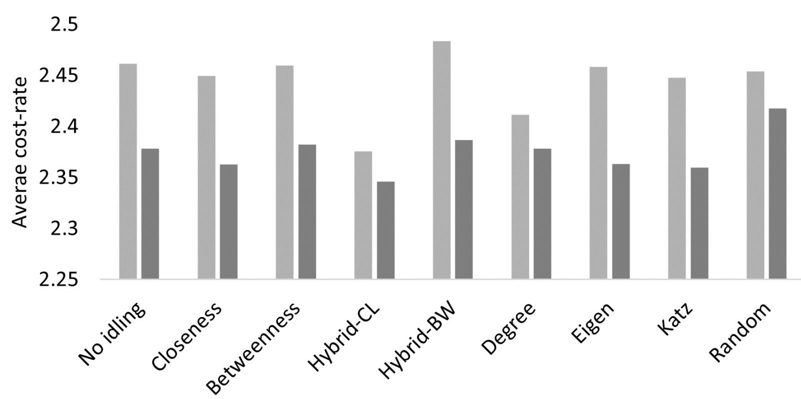
Figure 22: Mean cost-per-unit time under the heuristic policies across ten randomly generated *grid* graphs. In each chart, the value of a single parameter changes while all other parameter values remain equal to the benchmark scenario. Legend colors represent the policies as follows: Committed Maintenance, Committed Routing with Proximal Maintenance, and Dynamic Routing with Proximal Maintenance.



(a) Committed Maintenance



(b) Committed Routing with Proximal Maintenance



(c) Dynamic Routing with Proximal Maintenance

Figure 23: Heuristic policy performance under different approaches of obtaining the "anticipatory idling" set A for tree and grid graphs and the benchmark parameter setting.

Table 13: Percent decrease in the mean cost-per-unit time of Committed and Dynamic Routing with Proximal Maintenance policies by adopting repair threshold $\bar{\mu}^*$ obtained from equation (4.10). The percentages are averages across the ten random tree graph instances of Figure 21. A dynamic repair threshold significantly improves the performance of the Committed Routing with Proximal Maintenance policy.

		repair cost			downtime cost			travel cost			transition		
Policies with $\bar{\mu}^*$	benchmark	$c_R(0),\ldots,c_R(\Delta)$			c_D			c_T			matrix P		
obtained from (4.10)	scenario	60	100	160	100	120	160	1	5	10	P'	P''	P'''
Committed Routing	62%	53%	59%	46%	52%	31%	68%	29%	50%	48%	44%	28%	0%
Dynamic Routing	2%	-1%	0%	0%	0%	0%	0%	3%	1%	2%	2%	3%	1%

Table 14: Percent decrease in the mean cost-per-unit time of Committed and Dynamic Routing with Proximal Maintenance policies by adopting repair threshold $\bar{\mu}^*$ obtained from equation (4.10). The percentages are averages across the ten random grid graph instances of Figure 22. A dynamic repair threshold significantly improves the performance of the Committed Routing with Proximal Maintenance policy.

		repair cost			downtime cost			travel cost			transition		
Policies with $\bar{\mu}^*$	benchmark	$c_R(0),\ldots,c_R(\Delta)$			c_D			c_T			matrix P		
obtained from (4.10)	scenario	60	100	160	100	120	160	1	5	10	P'	P''	P'''
Committed Routing	53%	53 %	49%	53%	64%	38%	62%	64%	62%	48%	50%	35%	21%
Dynamic Routing	0%	0%	2%	2%	0%	1%	-1%	1%	1%	4%	0%	0%	0%

5.0 Summary of Contributions and Future Work

In Chapters 2 and 3 we relax the underlying assumptions of the age-replacement policy to ensure its practicability in general settings. Specifically, Chapter 2 generalizes previous work on unpunctual policy implementation by considering a setting where the degree of deviation from prescribed replacement times depends of the policy itself. We model the temporal impact of the prescribed replacement policy on the degree of deviation and analyze its effect on the optimal policy. We conclude that the optimal policy aims to decrease the expected degree of unpunctuality and its variance, and explore conditions under which the nonstationarity can be exploited to reduce the long-run cost-rate.

The insights generated by Chapter 2 apply to settings in which a maintenance planner prescribes an age-replacement policy with minimal repair. One may also consider unpunctual implementation for age-replacement *without* minimal repair. In the absence of minimal repair, the long-run cost rate minimizing replacement policy under punctual and stationary unpunctual behavior are modeled and studied in [13] and [64], respectively. In our modeling framework with non-stationary unpunctual behavior, the optimal age-replacement time without minimal repair corresponds to solving

$$\min_{T\in S}\Theta(T) = \frac{\int_a^b \Big(c_r F_X\big(T+yz(T)\big) + c_p \bar{F}_X\big(T+yz(T)\big)\Big) dF_Y(y)}{\int_a^b \int_0^{T+yz(T)} \bar{F}(x)dx dF_Y(y)}, \tag{5.1}$$

where c_r and c_p are the costs of reactive and preventive replacement, respectively. It could be interesting to explore whether the optimal replacement policy exhibits similar behavior to the results established in Chapter 2. As stated in [64], in the absence of minimal repair, it is more difficult to characterize how unpunctual policy implementation affects the optimal solution. Hence, analysis of (5.1) is likely to require novel analytical approaches or be largely numerical in nature.

Next, in Chapter 3, we relax the assumption of constant replacement costs for the age-replacement problem without minimal repair. In the literature on age-replacement policies with minimal repair, results have been provided that prove the existence and uniqueness of the optimal policy under various functional forms of age-dependent replacement costs. Yet, this body of literature does not compare optimal policies and cost-rates under different maintenance cost functions, or with policies under an availability-maximizing criterion. Extending and comparing our results to those under age-replacement policies with minimal repair could yield useful insights to maintenance planners.

Moreover, we restrict our analysis to cases under which replacement costs are either convex or concave. However, throughout system's lifetime, replacement cost functions may take both concave and convex forms (see e.g., Example 8) which could be because of a sudden decline in salvage value at some age(s). Such functional forms can result in near-optimal solution(s) that are far from the optimal solution, but may be more appealing from an implementation standpoint. Further exploring the long-run expected cost-rate under such step-like replacement costs may be an interesting direction for future research.

Age-dependent replacement costs have received little attention in the condition-based maintenance (CBM) literature. The study in [61] consider a condition based maintenance problem in which replacement costs depend on both system deterioration level and age, and provide a limited numerical study for this interesting setting. Considering age-dependent replacement costs in various CBM settings and exploring their effect on the maintenance thresholds can be an interesting direction for future research.

Finally, in Chapter 4, we integrate spatial considerations with maintenance planning for condition-based maintenance. The optimal policy derived from our Markov decision process formulation prescribes routing/positioning decisions of the maintenance resource and the timing of the interventions that the resource performs. We demonstrate these decisions are affected by a deteriorated asset's position on the graph and the overall graph structure. The insights developed from our theoretical derivations and numerical observations yield easy-to-implement yet high-performing heuristic policies that can be applied to large-scale applications.

Our study in Chapter 4 can be generalized in various directions for future research. Recall we assume that each repair action takes one unit of time, suggesting that our model captures settings where repair times are relatively shorter than traveling times. Applications with larger repair to travel time ratios may pose novel trade-offs not captured by our model. A subsequent study could explore how longer maintenance and downtime durations can affect the repositioning decisions and repair thresholds.

Another important direction is to consider multiple maintenance resources. In particular, one may formulate a mathematical program to optimally assign a resource to a *cluster* of assets and use our model as a sub-procedure for every resource-cluster pair. Such model can capture real-world practicalities where each resource is allocated to a particular geographical area. In such settings, high-performance heuristics may use graph clustering methods to identify the resource-cluster pairs using closeness-like centrality measures such as eccentricity, harmonic, decay, and their group centrality counterparts which identify the centrality of a group of nodes [129, 130].

Lastly, as stated earlier, our problem can suffer from curse of dimensionality for large instances. A common approach to overcome this challenge is to design basis functions to approximate the value function [91, 104]. We conjecture that incorporating centrality measures such as closeness centrality in basis functions can produce powerful approximations. Such parametric models can tackle the computational challenges of dynamic mobile maintenance resource problems.

Appendix A Chapter 2.0: Proofs

Proof of Proposition 1

First, note that $m(T)$ is continuous on $[0, +\infty)$. Then, the first derivative of $m(T)$ is given by:

$$
\begin{aligned}
m'(T) = {} & \frac{(1+\mu_Y z'(T))^2 - \mu_Y z''(T)(T+\mu_Y z(T))}{(1+\mu_Y z'(T))^2} \cdot \int_a^b (1+yz'(T))h_X(T+yz(T))dF_Y(y) \\
& + \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b yz''(T)h_X(T+yz(T))dF_Y(y) \\
& + \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 h'_X(T+yz(T))dF_Y(y) \\
& - \int_a^b (1+yz'(T))h_X(T+yz(T))dF_Y(y) \\
= {} & \frac{T+\mu_Y z(T)}{(1+\mu_Y z'(T))^2} \int_a^b z''(T)(y-\mu_Y)h_X(T+yz(T))dF_Y(y) \\
& + \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 h'_X(T+yz(T))dF_Y(y) \\
= {} & \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \cdot \frac{z''(T)}{1+\mu_Y z'(T)} \cdot \Big(\mathrm{E}[h_X(T+yz(T))\cdot Y] - \mathrm{E}[h_X(T+yz(T)] \cdot \mathrm{E}[Y]\Big) \\
& + \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 h'_X(T+yz(T))dF_Y(y) \\
= {} & \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \cdot \frac{z''(T)}{1+\mu_Y z'(T)} \mathrm{Cov}\Big(h_X(T+yz(T)), Y\Big) && \text{(A.1)} \\
& + \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 h'_X(T+yz(T))dF_Y(y). && \text{(A.2)}
\end{aligned}
$$

Next, because $z(T) \geq 0$ (Assumption 2), $h_X(T+yz(T))$ is strictly increasing in y. Hence, $\mathrm{Cov}\Big(h_X(T+yz(T)), Y\Big) \geq 0$. Therefore, if $1+\mu_Y z'(T) > 0$ and $z''(T) \geq 0$, then expression (A.1) is non-negative. Expression (A.2) is also non-negative by Assumption 1. As a result, $m(T)$ is increasing.

Next, it is sufficient to show that there exists a T' such that for all $T \geq T'$, $m'(T) \geq \epsilon > 0$ for some $\epsilon \in \mathbb{R}$. Given that $m(T)$ is increasing, the existence of such a T' implies that there exists a T such that $m(T) = k$. To establish this result, first note that by Assumption 1, there exists a strictly positive lower bound ℓ for $h'_X(T+yz(T))$. Thus, for sufficiently large values of T, we have

$$
\begin{aligned}
m'(T) &\geq \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 h'_X(T+yz(T)) f_Y(y) dy \\
&\geq \ell \cdot \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \int_a^b (1+yz'(T))^2 f_Y(y) dy \\
&= \ell \cdot \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \cdot \Big\{ \mathrm{Var}\big(1+Yz'(T)\big) + \Big[\int_a^b (1+yz'(T)) f_Y(y) dy \Big]^2 \Big\} \\
&= \ell \cdot \frac{T+\mu_Y z(T)}{1+\mu_Y z'(T)} \Big\{ \big(z'(T)\big)^2 \mathrm{Var}(Y) + \big(1+\mu_Y z'(T)\big)^2 \Big\} \\
&\geq \ell \cdot \big(T+\mu_Y z(T)\big)\big(1+\mu_Y z'(T)\big).
\end{aligned}
$$

Therefore, by condition (i) of Proposition 1,

$$
\lim_{T\to\infty} m'(T) \geq \big(T+\mu_Y z(T)\big)\delta\ell > 0,
$$

which implies the sufficient result.

Hence, $\Omega(T)$ is quasi-convex and if condition (2.9) holds, then there exists a unique solution T^* to problem (2.7). Otherwise, the minimum cost-rate is obtained by equation (2.10) because the minimum of $m(T)$ is achieved at time $T = \max\big\{\min\{t : t+az(t) > 0\}, 0\big\}$. $\square$

Proof of Theorem 1

See [13, p. 87].

Proof of Theorem 2

For the first inequality see [13, p. 87]. The second inequality holds because $\Omega(T^*) \leq \Omega(\hat{T}^*)$; and by equation (2.11), $\Omega(\hat{T}^*) = \hat{\Omega}(\hat{T}^*)$. $\square$

Proof of Theorem 3

Because T^* is sub-optimal for problem (2.6), $\hat{\Omega}(\hat{T}^*) \leq \hat{\Omega}(T^*)$. Moreover, because $z(T^*) = 1$, we have that $\hat{\Omega}(T^*) = \Omega(T^*)$. Therefore, $\hat{\Omega}(\hat{T}^*) \leq \Omega(T^*)$. $\square$

Proof of Theorem 4

Let $\mathcal{H}(u) = \int_0^u h(x)dx$. For $\mu_Y = 0$ and symmetric $f_Y(y)$ (i.e., $a = -b$), we have

$$
\Omega(T) - \hat{\Omega}(T) = \frac{\int_{-b}^{b} c_m \Big(\int_0^{T+yz(T)} h_X(x)dx - \int_0^{T+y} h_X(x)dx \Big) dF_Y(y)}{T}
$$

$$= \frac{c_m \int_{-b}^{b} \int_{T+y}^{T+yz(T)} h_X(x) dx dF_Y(y)}{T}$$

$$= \frac{c_m \int_{-b}^{b} \Big(\mathcal{H}(T+yz(T)) - \mathcal{H}(T+y) \Big) dF_Y(y)}{T}. \tag{A.3}$$

Now, let $\mathcal{G}(y,T) = \mathcal{H}(T+yz(T)) - \mathcal{H}(T+y)$. We let $\mathcal{G}(y,T) = \mathcal{G}_1(y,T) + \mathcal{G}_2(y,T)$ where $\mathcal{G}_1(y,T)$ and $\mathcal{G}_2(y,T)$ are given by:

$$\mathcal{G}_1(y,T) = \frac{\mathcal{G}(y,T) + \mathcal{G}(-y,T)}{2}$$

and

$$\mathcal{G}_2(y,T) = \frac{\mathcal{G}(y,T) - \mathcal{G}(-y,T)}{2}.$$

Then, equation (A.3) reduces to

$$\begin{aligned} \Omega(T) - \hat{\Omega}(T) =& \frac{c_m}{T} \int_{-b}^{b} \Big(\mathcal{G}_1(y,T) + \mathcal{G}_2(y,T) \Big) dF_Y(y) \\ =& \frac{2c_m}{T} \int_{0}^{b} \mathcal{G}_1(y,T) dF_Y(y), \end{aligned} \tag{A.4}$$

because $\int_{-b}^{b} \mathcal{G}_2(y,T) dF_Y(y) = 0$ for $\mu_Y = 0$ and symmetric $f_Y(y)$.

Observe that $\mathcal{H}(u)$ is both non-decreasing and convex because $h_X(x)$ is non-negative and increasing (Assumption 1). If $z(T) > 1$, then from equation (A.4) it suffices to show that $\mathcal{G}_1(y,T) > 0$, i.e.,

$$\begin{aligned} \mathcal{G}_1(y,T) =& \frac{\mathcal{H}(T+yz(T)) - \mathcal{H}(T+y) + \mathcal{H}(T-yz(T)) - \mathcal{H}(T-y)}{2} \\ =& \frac{\mathcal{H}(T+y+y(z(T)-1)) - \mathcal{H}(T+y)}{2} \\ &- \frac{\mathcal{H}(T-yz(T)+y(z(T)-1)) - \mathcal{H}(T-yz(T))}{2} \\ >& 0. \end{aligned} \tag{A.5}$$

Note that $y \in [0,b]$ and $z(T) > 1$. Thus, $T + yz(T) > T + y > T - y > T - yz(T)$. Moreover, $\big(T + yz(T)\big) - (T + y) = (T - y) - \big(T - yz(T)\big)$. Therefore, in equation (A.5), $\Big(\mathcal{H}(T+y+y(z(T)-1)) - \mathcal{H}(T+y)\Big) > \Big(\mathcal{H}(T-yz(T)+y(z(T)-1)) - \mathcal{H}(T-yz(T))\Big)$

because of increasing differences in the convex function $\mathcal{H}(u)$ (see Proposition 2.1.6 in [122, p. 20]). As a result, $\mathcal{G}_1(y,T) > 0$. By a similar argument, if $z(T) < 1$, then $\mathcal{G}_1(y,T) < 0$. The result follows directly. $\square$

Proof of Theorem 5

By Assumption 3 (condition (i) in Proposition 1) and $z''(T) \geq 0$ for all T (condition (ii) in Proposition 1), $m(T)$ is monotone increasing (recall the proof of Proposition 1). Moreover, in [64, see p. 140], it is shown that $\hat{m}(T)$ in equation (2.14) is also monotone increasing.

In order to compare T^* and $\hat{T}^*$ we compare $m(\hat{T}^*)$ and $\hat{m}(\hat{T}^*)$ (note that $\hat{m}(\hat{T}^*) = k$). If $m(\hat{T}^*) > k$, then $T^* < \hat{T}^*$. Conversely, if $m(\hat{T}^*) < k$, then $T^* > \hat{T}^*$. We have that

$$\begin{aligned}
m(\hat{T}^*) - \hat{m}(\hat{T}^*) &= \frac{\hat{T}^* + \mu_Y}{1 + \mu_Y z'(\hat{T}^*)} \int_a^b (1 + yz'(\hat{T}^*))h_X(\hat{T}^* + y)f_Y(y)dy \\
&\quad - (\hat{T}^* + \mu_Y) \int_a^b h_X(\hat{T}^* + y)f_Y(y)dy \\
&= \int_a^b \Big(\frac{1 + yz'(\hat{T}^*)}{1 + \mu_Y z'(\hat{T}^*)} - 1\Big)(\hat{T}^* + \mu_Y)h_X(\hat{T}^* + y)f_Y(y)dy \\
&= \int_a^b z'(\hat{T}^*) \cdot \frac{\hat{T}^* + \mu_Y}{1 + \mu_Y z'(\hat{T}^*)}(y - \mu_Y)h_X(\hat{T}^* + y)f_Y(y)dy \\
&= z'(\hat{T}^*) \cdot \frac{\hat{T}^* + \mu_Y}{1 + \mu_Y z'(\hat{T}^*)} \cdot \Big(\mathrm{E}[h_X(\hat{T}^* + y) \cdot Y] - \mathrm{E}[h_X(\hat{T}^* + y)] \cdot \mathrm{E}[Y]\Big) \\
&= z'(\hat{T}^*) \cdot \frac{\hat{T}^* + \mu_Y}{1 + \mu_Y z'(\hat{T}^*)} \cdot \mathrm{Cov}(h_X(\hat{T}^* + y), Y). \qquad \text{(A.6)}
\end{aligned}$$

Note that $\mathrm{Cov}(h_X(\hat{T}^* + y), Y) > 0$ because $h(\hat{T}^* + y)$ is strictly increasing in y. Also, $1 + \mu_Y z'(\hat{T}^*) > 0$ under Assumption 3. Thus, from equation (A.6) we conclude that if $z'(\hat{T}^*) > 0$, then $m(\hat{T}^*) > \hat{m}(\hat{T}^*)$. On the contrary, if $z'(\hat{T}^*) < 0$ then $m(\hat{T}^*) < \hat{m}(\hat{T}^*)$. The result follows directly. $\square$

Proof of Corollary 1

This result is proved by a similar approach used in the proof of Theorem 5.

Proof of Proposition 2

See Theorem 5 and Lemma 1.

Proof of Proposition 3

See Theorem 5 and Lemma 2.

Proof of Proposition 4

See Theorem 5 and Lemma 3.

Proof of Proposition 5

See Theorem 5 and Lemma 4.

Lemma 5 (Edmundson-Madansky inequality [68]). *Let $V \in [\dot{a}, \dot{b}]$ have a c.d.f. $F_V(v)$ and finite mean μ_V. Suppose ϕ is a bounded convex function of $v \in [\dot{a}, \dot{b}]$. An upper bound for* $\mathrm{E}[\phi(V)]$ *is*

$$\mathrm{E}[\phi(V)] \leq \frac{\phi(\dot{b}) - \phi(\dot{a})}{\dot{b} - \dot{a}}(\mu_V - \dot{a}) + \phi(\dot{a}).$$

Proof of Proposition 6

The first inequality in (2.15) follows directly from Theorem 2.

For the second inequality in (2.15), first note that given that $\widetilde{T}_1$ and $\widetilde{T}_2$ are feasible for problem (2.7), $\Omega(T^*) \leq \min\{\Omega(\widetilde{T}_1), \Omega(\widetilde{T}_2)\}$. Also, note that $\widetilde{\Omega}(\widetilde{T}^*) = c_m h_X(\widetilde{T}^*)$ [13]. Thus, U_Y^1 is obtained by dividing $\Omega(T)$ in equation (2.7) by $c_m h_X(\widetilde{T}^*)$.

For the third inequality in (2.15), first note that $\phi(y) = \int_0^{T+yz(T)} h_X(x)dx$ is a bounded convex function of y because $\phi''(y) = \big(z(T)\big)^2 h'_X\big(T + yz(T)\big) \geq 0$ (Assumption 1). Hence, we can apply the Edmundson-Madansky inequality in Lemma 5 and for $Y \in [a, b]$ we obtain

$$\begin{aligned} M(T) =& \mathrm{E}\Big[\int_0^{T+yz(T)} h_X(x)dx\Big] \\ \leq& \frac{\int_{T+az(T)}^{T+bz(T)} h_X(x)dx}{b-a}(\mu_Y - a) + \int_0^{T+az(T)} h_X(x)dx. \end{aligned} \tag{A.7}$$

The upper bound $U^1(Y)$ in Proposition 6 is obtained accordingly. □

Proof of Proposition 7

For the first inequality in (2.16), first note that from Theorem 2, $\widetilde{\Omega}(\widetilde{T}^*) \leq \Omega(T^*)$. Next, observe that given that $\widetilde{T}^* - \mu_Y$ is feasible for problem (2.6), $\hat{\Omega}(\widetilde{T}^* - \mu_Y) \geq \hat{\Omega}(\hat{T}^*)$. Therefore, the lower bound in (2.16) is obtained by dividing $\widetilde{\Omega}(\widetilde{T}^*) = c_m h_X(\widetilde{T}^*)$ by $\hat{\Omega}(\widetilde{T}^* - \mu_Y)$.

The second inequality in (2.16) directly follows from Theorem 2. □

Proof of Proposition 8

For the first inequality in (2.17), L_Y^2 is obtained by dividing $\Omega(\widetilde{T}^*)$ by $\widetilde{\Omega}(T)$. The result holds because $\widetilde{T}_1$ and $\widetilde{T}_2$ are sub-optimal for problem (2.7).

For the second inequality, recall that $\widetilde{\Omega}(\widetilde{T}^*) = c_m h_X(\widetilde{T}^*)$, and $\widetilde{\Omega}(\widetilde{T}^*) \leq \Omega(T^*)$ by Theorem 2. Thus, $U_Y^2(\widetilde{T}^*)$ is obtained by dividing $\Omega(\widetilde{T}^*)$ by $c_m h_X(\widetilde{T}^*)$.

Similar to the proof of Proposition 6, we let $\phi(y) = \int_0^{T+yz(T)} h_X(x)dx$, and note that $\phi(y)$ is a bounded convex function of y. Hence, for the third inequality we apply the Edmundson-Madansky inequality in Lemma 5 and obtain $M(T)$ as in equation (A.7). The upper bound $U^2(Y)$ in Proposition 8 is obtained accordingly. $\square$

Appendix B Chapter 3.0: Proofs

Additional Notation and Previous Results

We take the approach in [13, p. 85]. A necessary and sufficient condition that T^* minimizes the cost-rate function in (3.2) is obtained by setting its derivative equal to zero:

$$m(T) \equiv \Big\{\big(c_r(T) - c_p(T)\big)h(T) + c_p'(T)\Big\} \int_0^T \bar{F}(t)dt - \int_0^T c_r(t)f(t)dt - c_p(T)\bar{F}(T) = 0. \tag{B.1}$$

That is, if equation (B.1) has a unique solution, then that solution minimizes the cost-rate in problem (3.2); i.e., the solution to (B.1) is T^*.

Next, we provide proofs for existence and uniqueness of $\widetilde{T}^*$ and $T^*_{\mathcal{A}}$. Recall that $\widetilde{T}^*$ denotes the optimal policy to minimize the long-run expected cost-rate under constant replacement costs (see Section 3.4), and $T^*_{\mathcal{A}}$ denotes the optimal policy to maximize availability (see Section 3.5).

Existence and Uniqueness of $\widetilde{T}^*$ (see [13, p. 87])

By setting the first derivative of the objective function in problem (3.3) equal to zero, we have

$$\widetilde{m}(T) \equiv (\widetilde{c}_r - \widetilde{c}_p)h(T) \int_0^T \bar{F}(t)dt - \widetilde{c}_r F(T) - \widetilde{c}_p \bar{F}(T) = 0. \tag{B.2}$$

It has been previously shown in [13, see p. 87] that $\widetilde{m}(T)$ is monotone increasing, and as a result $\widetilde{\Omega}(T)$ is quasi-convex. By Assumption 4, $h(T)$ is strictly increasing and hence, $\lim_{T\to\infty} \widetilde{m}(T) > 0$. Finally, because $m(0) < 0$, there exists a unique and finite optimal solution to equation (B.2).

Existence and Uniqueness of $T^*_{\mathcal{A}}$ (see [96, p. 139]))

By setting the first derivative of the objective function in problem (3.5) equal to zero, we have

$$m_{\mathcal{A}}(T) \equiv (\beta_r - \beta_p)F(T) + \beta_p - (\beta_r - \beta_p)h(T) \int_0^T \bar{F}(t)dt = 0. \tag{B.3}$$

The function $m_{\mathcal{A}}(T)$ is monotone decreasing, and as a result $\widetilde{\Omega}(T)$ is quasi-concave. If $\beta_r > \beta_p$, then by Assumption 4, $h(T)$ is strictly increasing and hence, $\lim_{T\to\infty} m_{\mathcal{A}}(T) < 0$. Finally, because $m_{\mathcal{A}}(0) > 0$, there exists a unique and finite optimal solution to equation (B.3). If $\beta_r = \beta_p = \beta$, then $m_{\mathcal{A}}(T) = \beta$. Hence, $\mathcal{A}(T)$ is strictly increasing from 0 to $\mu/(\mu+\beta)$ and $T^*_{\mathcal{A}} = +\infty$.

Proof of Theorem 6

Recall that if T^* exists, then it is the solution to equation (B.1) (see Appendix A). First, note that $m(0) = -c_p(0) < 0$ (because $F(0) = 0$). Hence, to ensure equation (B.1) has exactly one solution, it suffices to show that $m(T)$ is strictly increasing to infinity. We take the first derivative of $m(T)$:

$$m'(T) = \Big\{\big(c_r(T) - c_p(T)\big)h'(T) + \big(c_r'(T) - c_p'(T)\big)h(T) + c_p''(T)\Big\} \int_0^T \bar{F}(t)dt. \tag{B.4}$$

Now, we prove the uniqueness result under conditions (i) and (ii). By equation (B.4) as well as Assumptions 4 and 5, if $c_r'(T) \geq c_p'(T)$ and $c_p''(T) \geq 0$ for all T, then $m'(T) > 0$. Because $m(T)$ is increasing, the cost-rate function $\Omega(T)$ is quasi-convex and has at most one solution.

Next, to ensure that the unique optimal solution is finite, it is sufficient to show that $m(T)$ increases to infinity as T increases. Hence, we show that $\lim_{T\to\infty} m'(T) > \epsilon$ for some $\epsilon > 0$. By Assumptions 4 and 5, we have

$$\lim_{T\to\infty} m'(T) \geq \big(c_r(\infty) - c_p(\infty)\big)h'(\infty)\cdot\mu = \delta h'(\infty)\mu = \epsilon > 0, \tag{B.5}$$

which implies the sufficient result that the unique optimal solution is finite.

Next, we prove the uniqueness result under conditions $(ii) - (v)$. Similar to the previous part, it suffices to show $m'(T)$ is non-negative and $\lim_{T\to\infty} m'(T) > \epsilon$ for some $\epsilon > 0$. From equation (B.4) we have

$$m'(T) = \mathcal{G}(T)\int_0^T \bar{F}(t)dt,$$

where

$$\mathcal{G}(T) = \big(c_r(T) - c_p(T)\big)h'(T) + \big(c_r'(T) - c_p'(T)\big)h(T) + c_p''(T).$$

First, note that $\int_0^T \bar{F}(t)dt$ is an increasing, non-negative function which equals 0 at time 0 and converges to μ. Next, note that

$$\mathcal{G}(0) = \big(c_r(0) - c_p(0)\big)h'(0) + c_p''(0);$$

thus, we can have either $\mathcal{G}(0) \geq 0$ or $\mathcal{G}(0) \leq 0$. We take the first derivative of $\mathcal{G}(T)$:

$$\mathcal{G}'(T) = \big(c_r(T) - c_p(T)\big)h''(T) + 2\big(c_r'(T) - c_p'(T)\big)h'(T) + \big(c_r''(T) - c_p''(T)\big)h(T) + c_p'''(T). \quad \text{(B.6)}$$

From equation (B.6) as well as conditions $(ii) - (v)$, $\mathcal{G}'(T)$ is non-negative. As a result, if $\mathcal{G}(0) \geq 0$, then $\mathcal{G}(T) \geq 0$ and consequently, $m'(T) \geq 0$, which implies the uniqueness result. On the other hand, when $\mathcal{G}(0) < 0$, because $\mathcal{G}(T)$ is increasing, $m'(T)$ first takes negative values for some $T < T^*_{\mathcal{G}}$, where $T^*_{\mathcal{G}}$ is the solution to $\mathcal{G}(T) = 0$, and then takes positive values. As a result, when $\mathcal{G}(0) \leq 0$, the function $m(T)$ has only one local minimum at $T^*_{\mathcal{G}}$ with $m(T^*_{\mathcal{G}}) < 0$. Then, equation (B.1) has at most one solution.

Finally, by Assumption 5 and $h''(T) > 0$, we have $\lim_{T\to\infty} \mathcal{G}'(T) > \epsilon_1$ for some $\epsilon_1 > 0$. Hence, $\lim_{T\to\infty} m'(T) > \epsilon$ for some $\epsilon > 0$ and there exists an optimal solution to equation (B.1). As a direct result of equation (B.1), we have that $\Omega(T^*) = \big(c_r(T^*) - c_p(T^*)\big)h(T^*) + c_p'(T^*)$. □

Proof of Theorem 7

Recall from the proof of Theorem 6 that by Assumption 6, $m(T)$ is monotone increasing for all $T > 0$ or $T > T_{\mathcal{G}}$, where $T_{\mathcal{G}}$ is unique with $m'(T_{\mathcal{G}}) = 0$ and $m(T_{\mathcal{G}}) < 0$. Furthermore, $\widetilde{m}(T)$ is monotone increasing [13, p. 85]. Therefore, in order to compare T^* and $\widetilde{T}^*$, we compare $m(\widetilde{T}^*)$ and $\widetilde{m}(\widetilde{T}^*)$. Let $\bar{m}(T) = m(T) - \widetilde{m}(T)$. From equations (B.1), (B.2), $c_r(t) = \widetilde{c}_r + g(t)$ and $c_p(t) = \widetilde{c}_p + g(t)$ we have

$$\begin{aligned}\bar{m}(T) =& g'(T)\int_0^T \bar{F}(t)dt - \int_0^T c_r(t)f(t)dt - c_p(T)\bar{F}(T) + \widetilde{c}_r F(T) + \widetilde{c}_p\bar{F}(T)\\ =& g'(T)\int_0^T \bar{F}(t)dt - \int_0^T g(t)f(t)dt - g(T)\bar{F}(T).\end{aligned} \quad \text{(B.7)}$$

From equation (B.7), $\bar{m}(0) = -g(0) = 0$ and $\bar{m}'(T) = g''(T)\int_0^T \bar{F}(t)dt$. Consequently, if $g''(T) \geq 0$, then $\bar{m}(T) \geq 0$ and $m(\widetilde{T}^*) \geq \widetilde{m}(\widetilde{T}^*)$. Conversely, if $g''(T) \leq 0$, then $\bar{m}(T) \leq 0$ and $m(\widetilde{T}^*) \leq \widetilde{m}(\widetilde{T}^*)$. Lastly, if $g''(T) = 0$, then $\bar{m}(T) = 0$ and $m(\widetilde{T}^*) = \widetilde{m}(\widetilde{T}^*)$. The results follow directly.

To finish the proof, we have

$$\Omega(T) - \widetilde{\Omega}(T) = \frac{\int_0^T g(x)f(x)dx + g(T)\bar{F}(T)}{\int_0^T \bar{F}(x)dx}. \quad \text{(B.8)}$$

The right-hand side in (B.8) is non-negative for all T. Thus, $\Omega(T^*) \geq \widetilde{\Omega}(\widetilde{T}^*)$. For the case when $g(T) = aT$, recall that we have $T^* = \widetilde{T}^*$ by the first part of the proof. Also, recall than $\Omega(T^*) = \big(c_r(T^*) - c_p(T^*)\big)h(T^*) + c_p'(T^*)$ (see Theorem 6). Then, $\Omega(T^*) - \widetilde{\Omega}(\widetilde{T}^*) = \Omega(T^*) - \widetilde{\Omega}(T^*) = c_p'(T^*) = a$. $\square$

Proof of Proposition 9

We take a similar approach as in the proof of Theorem 7 and compare functions (B.1) and (B.2) at time $\widetilde{T}^*$. From equations (B.1) and (B.2) we have

$$m(\widetilde{T}^*) - \widetilde{m}(\widetilde{T}^*) = \Big(c_r(\widetilde{T}^*) - \widetilde{c}_r\Big)h(\widetilde{T}^*)\int_0^{\widetilde{T}^*} \bar{F}(t)dt - \int_0^{\widetilde{T}^*} \Big(c_r(t) - \widetilde{c}_r\Big)f(t)dt \tag{B.9}$$

$$-\Big(c_p(\widetilde{T}^*) - \widetilde{c}_p\Big)h(\widetilde{T}^*)\int_0^{\widetilde{T}^*} \bar{F}(t)dt + c_p'(\widetilde{T}^*)\int_0^{\widetilde{T}^*} \bar{F}(t)dt - \Big(c_p(\widetilde{T}^*) - \widetilde{c}_p\Big)\bar{F}(\widetilde{T}^*). \tag{B.10}$$

By Assumption 5, if $c_p(\widetilde{T}^*) \leq \widetilde{c}_p$, then the term in (B.10) is non-negative. From the term in (B.9), let

$$\mathcal{H}(T) = \Big(c_r(T) - \widetilde{c}_r\Big)h(T)\int_0^T \bar{F}(t)dt - \int_0^T \Big(c_r(t) - \widetilde{c}_r\Big)f(t)dt.$$

First, note that $\mathcal{H}(0) = 0$. Next, we have

$$\mathcal{H}'(T) = c_r'(T)h(T)\int_0^T \bar{F}(t)dt + \Big(c_r(T) - \widetilde{c}_r\Big)h'(T)\int_0^T \bar{F}(t)dt.$$

Given Assumptions 4 and 5, if $c_r(T) \geq \widetilde{c}_r$ for all $T \leq \widetilde{T}^*$, then $\mathcal{H}'(T)$ is non-negative for all $T \leq \widetilde{T}^*$. Hence, $\mathcal{H}(T) \geq 0$ for all $T \leq \widetilde{T}^*$. Consequently, $m(\widetilde{T}^*) \geq \widetilde{m}(\widetilde{T}^*)$ and $T^* \leq \widetilde{T}^*$. $\square$

Proof of Theorem 8

We take the same approach in the proof of Proposition 9 and compare $m(\widetilde{T}^*)$ and $\widetilde{m}(\widetilde{T}^*)$:

$$\begin{aligned}
m(\widetilde{T}^*) - \widetilde{m}(\widetilde{T}^*) =& \Big\{\big(\widetilde{c}_r - \widetilde{c}_p\big)h(\widetilde{T}^*) + c_p'(\widetilde{T}^*)\Big\}\int_0^{\widetilde{T}^*} \bar{F}(t)dt - \int_0^{\widetilde{T}^*} c_r(t)f(t)dt - \widetilde{c}_p\bar{F}(\widetilde{T}^*) \\
& - (\widetilde{c}_r - \widetilde{c}_p)h(\widetilde{T}^*)\int_0^{\widetilde{T}^*} \bar{F}(t)dt + \widetilde{c}_r F(\widetilde{T}^*) + \widetilde{c}_p\bar{F}(\widetilde{T}^*) \\
=& c_p'(\widetilde{T}^*)\int_0^{\widetilde{T}^*} \bar{F}(t)dt + \int_0^{\widetilde{T}^*} \Big(\widetilde{c}_r - c_r(t)\Big)\bar{f}(t)dt.
\end{aligned} \tag{B.11}$$

By Assumption 5, $c_r(T) \leq \widetilde{c}_r$ for all $T \leq \widetilde{T}^*$ and $c_p'(\widetilde{T}^*) \geq 0$. Thus, by equation (B.11), $m(\widetilde{T}^*) \geq \widetilde{m}(\widetilde{T}^*)$ and consequently, $T \leq \widetilde{T}^*$.

Next, we show that $\Omega(T^*) < \widetilde{\Omega}(\widetilde{T}^*)$. Note that

$$\Omega(T^*) < \Omega(\widetilde{T}^*) = \frac{\int_0^{\widetilde{T}^*} c_r(t)f(t)dt + c_p(\widetilde{T}^*)\bar{F}(\widetilde{T}^*)}{\int_0^{\widetilde{T}^*} \bar{F}(t)dt}. \tag{B.12}$$

Thus, the inequality

$$\begin{aligned}
\Omega(\widetilde{T}^*) - \widetilde{\Omega}(\widetilde{T}^*) =& \frac{\int_0^{\widetilde{T}^*} c_r(t)f(t)dt + c_p(\widetilde{T}^*)\bar{F}(\widetilde{T}^*)}{\int_0^{\widetilde{T}^*} \bar{F}(t)dt} - \frac{\widetilde{c}_r F(\widetilde{T}^*) + \widetilde{c}_p \bar{F}(\widetilde{T}^*)}{\int_0^{\widetilde{T}^*} \bar{F}(t)dt} \\
=& \frac{\int_0^{\widetilde{T}^*} \big(c_r(t) - \widetilde{c}_r\big)f(t)dt + \big(c_p(\widetilde{T}^*) - \widetilde{c}_p\big)\bar{F}(\widetilde{T}^*)}{\int_0^{\widetilde{T}^*} \bar{F}(t)dt} < 0
\end{aligned}$$

holds because by Assumptions 5, $c_r(T) \leq \widetilde{c}_r$ for all $T \leq \widetilde{T}^*$ and $c_p(\widetilde{T}^*) = \widetilde{c}_p$. Consequently, by inequality (B.12) we have that $\Omega(T^*) < \widetilde{\Omega}(\widetilde{T}^*)$. □

Proof of Theorem 9

Note that if $\beta_r = \beta_p = \beta$, then

$$\min_{T>0} \mathcal{C}(T) = \frac{\int_0^T c_r(x)f(x)dx + c_p(T)\bar{F}(T)}{\int_0^T \bar{F}(x)dx + \beta}.$$

Hence, the proof of the first statement in Theorem 9 is very similar to that of Theorem 6, and it is omitted. The second statement follows directly from the fact that $T^*_{\mathcal{C}}$ is finite, while $T^*_{\mathcal{A}} = +\infty$. □

Appendix C Chapter 4.0: Proofs

Additional Notation and Lemma

For the deterioration vector $\boldsymbol{x} = (x_1, \ldots, x_i, \ldots, x_{n_M})$, define $\boldsymbol{x}^{(i)+1} = (x_1^{(i)+1}, \ldots, x_i^{(i)+1}, \ldots, x_{n_M}^{(i)+1}) = (x_1, \ldots, x_i + 1, \ldots, x_{n_M})$. That is, the deterioration condition of the ith asset in $\boldsymbol{x}^{(i)+1}$ is one unit larger than that in $\boldsymbol{x}$, and all other conditions remain unchanged. Next, we establish Lemma 6.

Lemma 6. *If P has the IFR property and $v(\boldsymbol{x}', l) = v\big((x'_1, ..., x'_i, ..., x'_{n_M}), l\big)$ is nondecreasing in x'_i, then*

$$\sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', l) \leq \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j^{(i)+1}, x'_j} \cdot v(\boldsymbol{x}', l).$$

Proof. By the IFR property, for all $k \in \mathcal{K}$ we have

$$\sum_{x'_i = k}^{\Delta} P_{x_i, x'_i} \leq \sum_{x'_i = k}^{\Delta} P_{x_i+1, x'_i} = \sum_{x'_i = k}^{\Delta} P_{x_i^{(i)+1}, x'_i}. \tag{C.1}$$

Then, from equation (C.1) and that $v(\boldsymbol{x}', l)$ is nondecreasing in x'_i we have

$$\sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} P_{x_i, x'_i} \cdot \Big(\prod_{j \in V_M \setminus \{i\}} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', l) \Big) \leq \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} P_{x_i^{(i)+1}, x'_i} \cdot \Big(\prod_{j \in V_M \setminus \{i\}} P_{x_j, x'_j} \cdot v(\boldsymbol{x}', l) \Big) \tag{C.2}$$

$$= \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} P_{x_i^{(i)+1}, x'_i} \cdot \Big(\prod_{j \in V_M \setminus \{i\}} P_{x_j^{(i)+1}, x'_j} \cdot v(\boldsymbol{x}', l) \Big). \tag{C.3}$$

Inequality (C.2) is a result of Lemma 4.7.2. in [106, see p. 106], and equality (C.3) holds by the definition of $\boldsymbol{x}^{(i)+1}$. The result follows directly. □

Proof of Proposition 10

We construct the proof by induction on the iterates of the value iteration algorithm described in [106, see p. 161]. Let $v^k(\boldsymbol{x}, l)$ represent the cost obtained in the kth iteration of the algorithm. Because the value iteration algorithm converges for any set of initial values, without loss of generality we assume that

$$v^0(\boldsymbol{x}, l) = v_0, \text{ for all } (\boldsymbol{x}, l) \in S.$$

It follows directly that for all $(\boldsymbol{x}, l) \in S$, $v^0\big((x_1, \ldots, x_i, \ldots, x_{n_M}), l\big)$ is nondecreasing in x_i with all other state variables fixed. We now assume that for all $(\boldsymbol{x}, l) \in S$, $v^k\big((x_1, \ldots, x_i, \ldots, x_{n_M}), l\big)$ is also nondecreasing in x_i with all other state variables fixed. Thus, to complete the proof of Proposition 10 we need to show that $v^{k+1}\big((x_1, \ldots, x_i, \ldots, x_{n_M}), l\big)$ is nondecreasing in x_i with all other state variables fixed.

For now, we assume that $l \in V_M$ so that it is feasible to choose the repair action. Recall from Section 4.2 that $c_R(x_l)$ is nondecreasing in x_l. Then, from Lemma 6 we have

$$
\begin{aligned}
& v^{k+1}(\boldsymbol{x}, l) \\
=& \min \left\{ \begin{array}{l}
\mathcal{R}(\boldsymbol{x}, l) \equiv c_R(x_l) + c_D + \displaystyle\sum_{j \in V_M \setminus \{l\}} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}' \in \mathcal{K}^{n_M} \\ \text{s.t. } x'_l = 0}} \prod_{j \in V_M \setminus \{l\}} P_{x_j, x'_j} \cdot v^k(\boldsymbol{x}', l), \\
\mathcal{DN}(\boldsymbol{x}, l) \equiv \displaystyle\sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j, x'_j} \cdot v^k(\boldsymbol{x}', l), \\
\displaystyle\min_{b:(l,b) \in E} \mathcal{T}_b(\boldsymbol{x}, l) \equiv c_T + \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \min_{b:(l,b) \in E} \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j, x'_j} \cdot v^k(\boldsymbol{x}', b)
\end{array} \right\} \\
\leq& \min \left\{ \begin{array}{l}
\mathcal{R}(\boldsymbol{x}^{(i)+1}, l) \equiv \\
\quad c_R(x_l^{(i)+1}) + c_D + \displaystyle\sum_{j \in V_M \setminus \{l\}} c_D \cdot \mathbf{1}_{\{x_j^{(i)+1} = \Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}' \in \mathcal{K}^{n_M} \\ \text{s.t. } x'_l = 0}} \prod_{j \in V_M \setminus \{l\}} P_{x_j^{(i)+1}, x'_j} \cdot v^k(\boldsymbol{x}', l), \\
\mathcal{DN}(\boldsymbol{x}^{(i)+1}, l) \equiv \displaystyle\sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j^{(i)+1} = \Delta\}} + \lambda \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j^{(i)+1}, x'_j} \cdot v^k(\boldsymbol{x}', l), \\
\displaystyle\min_{b:(l,b) \in E} \mathcal{T}_b(\boldsymbol{x}^{(i)+1}, l) \equiv \\
\quad c_T + \displaystyle\sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j^{(i)+1} = \Delta\}} + \lambda \min_{b:(l,b) \in E} \sum_{\boldsymbol{x}' \in \mathcal{K}^{n_M}} \prod_{j \in V_M} P_{x_j^{(i)+1}, x'_j} \cdot v^k(\boldsymbol{x}', b)
\end{array} \right\} \qquad \text{(C.4)} \\
=& v^{k+1}(\boldsymbol{x}^{(i)+1}, l).
\end{aligned}
$$

Inequality (C.4) holds because $\mathcal{R}(\boldsymbol{x}, l) \leq \mathcal{R}(\boldsymbol{x}^{(i)+1}, l)$, $\mathcal{DN}(\boldsymbol{x}, l) \leq \mathcal{DN}(\boldsymbol{x}^{(i)+1}, l)$, and $\mathcal{T}_b(\boldsymbol{x}, l) \leq \mathcal{T}_b(\boldsymbol{x}^{(i)+1}, l)$ for all b such that $(l, b) \in E$. Hence, the induction hypothesis holds for $n \in \{0, 1, \ldots, k+1\}$. The proof follows similarly when $l \notin V_M$ and only the do nothing and travel actions are allowed, or when $i = l$. $\square$

Proof of Theorem 10

Proof under conditions of (i). We need to show that if $\mathcal{R}(\boldsymbol{x}, i) < \mathcal{DN}(\boldsymbol{x}, i)$ and $\mathcal{R}(\boldsymbol{x}, i) < \min_{b:(l,b)\in E} \mathcal{T}_b(\boldsymbol{x}, i)$, then $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) < \mathcal{DN}(\boldsymbol{x}^{(i)+1}, i)$ and $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) < \min_{b:(l,b)\in E} \mathcal{T}_b(\boldsymbol{x}^{(i)+1}, i)$. First, assume for contradiction that $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) < \mathcal{DN}(\boldsymbol{x}^{(i)+1}, i)$ does not hold. Hence, we have the following inequalities from $\mathcal{R}(\boldsymbol{x}, i) < \mathcal{DN}(\boldsymbol{x}, i)$ and $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) \geq \mathcal{DN}(\boldsymbol{x}^{(i)+1}, i)$, respectively:

$$c_R(x_i) + c_D + \sum_{j\in V_M\setminus\{i\}} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}'\in\mathcal{K}^{n_M} \\ \text{s.t. } x_i'=0}} \prod_{j\in V_M\setminus\{i\}} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i) < \sum_{j\in V_M} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i), \tag{C.5}$$

and

$$c_R(x_i+1) + c_D + \sum_{j\in V_M\setminus\{i\}} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}'\in\mathcal{K}^{n_M} \\ \text{s.t. } x_i'=0}} \prod_{j\in V_M\setminus\{i\}} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i) \geq \sum_{j\in V_M} c_D \cdot \mathbf{1}_{\{x_j^{(i)+1}=\Delta\}} + \lambda \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j^{(i)+1},x_j'} \cdot v(\boldsymbol{x}', i). \tag{C.6}$$

Subtracting equation (C.5) from (C.6), yields

$$c_R(x_i+1) - c_R(x_i) > c_D \cdot \mathbf{1}_{\{x_i+1=\Delta\}} + \lambda \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \left\{ \prod_{j\in V_M} P_{x_j^{(i)+1},x_j'} \cdot v(\boldsymbol{x}', i) - \prod_{j\in V_M} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i) \right\}. \tag{C.7}$$

The right-hand side of inequality (C.7) is non-negative by Lemma 6. Moreover, the left-hand side of (C.7) is 0 because the repair cost is constant under the set of conditions delineated in (i). Hence, inequality (C.6) cannot hold.

Second, assume for contradiction that $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) < \min_{b:(l,b)\in E} \mathcal{T}_b(\boldsymbol{x}^{(i)+1}, i)$ does not hold. Hence, we have the following inequalities from $\mathcal{R}(\boldsymbol{x}, i) < \min_{b:(l,b)\in E} \mathcal{T}_b(\boldsymbol{x}, i)$ and $\mathcal{R}(\boldsymbol{x}^{(i)+1}, i) \geq \min_{b:(l,b)\in E} \mathcal{T}_b(\boldsymbol{x}^{(i)+1}, i)$, respectively:

$$\begin{aligned} c_R(x_i) + c_D + \sum_{j\in V_M\setminus\{i\}} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}'\in\mathcal{K}^{n_M} \\ \text{s.t. } x_i'=0}} \prod_{j\in V_M\setminus\{i\}} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i) < \\ c_T + \sum_{j\in V_M} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \min_{b:(l,b)\in E} \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', b), \end{aligned} \tag{C.8}$$

and

$$\begin{aligned} c_R(x_i+1) + c_D + \sum_{j\in V_M\setminus\{i\}} c_D \cdot \mathbf{1}_{\{x_j=\Delta\}} + \lambda \sum_{\substack{\boldsymbol{x}'\in\mathcal{K}^{n_M} \\ \text{s.t. } x_i'=0}} \prod_{j\in V_M\setminus\{i\}} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', i) \geq \\ c_T + \sum_{j\in V_M} c_D \cdot \mathbf{1}_{\{x_j^{(i)+1}=\Delta\}} + \lambda \min_{b:(l,b)\in E} \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j^{(i)+1},x_j'} \cdot v(\boldsymbol{x}', b). \end{aligned} \tag{C.9}$$

Subtracting equation (C.8) from (C.9), yields

$$\begin{aligned} & c_R(x_i+1) - c_R(x_i) > c_D \cdot \mathbf{1}_{\{x_i+1=\Delta\}} \\ & \quad + \lambda \left\{ \min_{b:(l,b)\in E} \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j^{(i)+1},x_j'} \cdot v(\boldsymbol{x}', b) - \min_{b:(l,b)\in E} \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', b) \right\}. \end{aligned} \tag{C.10}$$

By Lemma 6, for any b we have

$$\sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j^{(i)+1},x_j'} \cdot v(\boldsymbol{x}', b) \geq \sum_{\boldsymbol{x}'\in\mathcal{K}^{n_M}} \prod_{j\in V_M} P_{x_j,x_j'} \cdot v(\boldsymbol{x}', b),$$

and thus the right-hand side of inequality (C.10) is non-negative. Moreover, the left-hand side of (C.10) is 0 because the repair cost is constant under the set of conditions delineated in (i). Hence, inequality (C.9) cannot hold.

Proof under conditions of (ii). The steps for the proof under conditions of (ii) are very similar to those outlined above. Note inequalities (C.7) and (C.10) with non-negative right-hand sides. If $x_i < \Delta - 1$, then the left-hand sides of (C.7) and (C.10) are non-positive because the repair cost is assumed to be constant for $x_i < \Delta$. On the other hand, if $x_i = \Delta - 1$, by the conditions delineated in (ii) we have $c_R(x_i+1) - c_R(x_i) \leq c_D$. Whereas, by (C.7) and (C.10) we have $c_R(x_i+1) - c_R(x_i) > c_D$. Hence, inequalities (C.6) and (C.9) cannot hold. □

Proof of Theorem 11

We have

$$v(\boldsymbol{x}, l) = \mathcal{T}_b(\boldsymbol{x}, l) = c_T + \sum_{j \in V_M} c_D \cdot \mathbf{1}_{\{x_j = \Delta\}} + \lambda \mathbb{E}\big[v(\boldsymbol{x'}, b) | \boldsymbol{X} = \boldsymbol{x}\big] \tag{C.11}$$

$$\geq \lambda \mathbb{E}\big[v(\boldsymbol{x'}, b) | \boldsymbol{X} = \boldsymbol{x}\big] \tag{C.12}$$

$$\geq \lambda v(\boldsymbol{x}, b), \tag{C.13}$$

where $\boldsymbol{X}$ is the vector of random variables denoting asset deterioration conditions. Equation (C.11) holds by definition (see equation (4.7)), and inequality (C.12) holds by the non-negativity of travel and downtime costs. Next, with regard to (C.13), note that under the assumptions of Theorem 11, transition probability matrix P has the IFR and upper triangular properties. The upper triangular property ensures that assets transition to deterioration levels that are greater than or equal to the current levels. Moreover, the IFR property is a sufficient condition for the result established in Proposition 10. Consequently, inequality (C.13) holds by Proposition 10. The first result of Theorem 11 follows directly.

For the second result, we first establish the proof under conditions of (i) followed by that under conditions of (ii).

Proof under conditions of (i). Assume for contradiction that if $v(\boldsymbol{x}, l) < \lambda v(\boldsymbol{x}, b)$, then $v(\boldsymbol{x}, l) \geq \mathcal{T}_b(\boldsymbol{x}, l)$. Note that by (4.4), the latter inequality holds by equality. That is,

$$v(\boldsymbol{x}, l) = \mathcal{T}_b(\boldsymbol{x}, l) < \lambda v(\boldsymbol{x}, b). \tag{C.14}$$

Moreover, The result follows because inequality (C.14) contradicts (C.13).

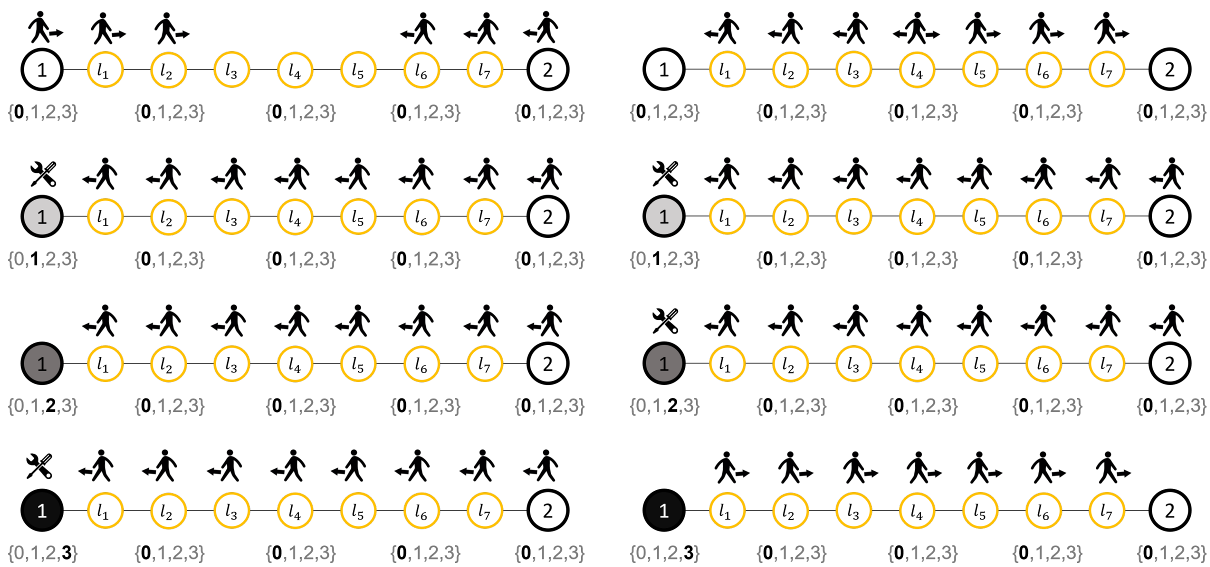
Proof under conditions of (ii). The steps of the proof are very similar to those under conditions of (i) with the exception that the inequality in (C.14) is not strict whereas the inequalities in (C.12) and (C.13) are strict because $c_T > 0$. □

Appendix D Chapter 4.0: Control-Limit Rule Violation

Here we present two numerical examples in which the conditions of Theorem 10, and thus the control-limit structure, are violated. These resulting optimal policies are depicted in Figure 24. Let $c_R(0) = 0$, $c_R(1) = 0$, $c_R(2) = 40$, $c_D = 41$, $c_T = 0.5$, and

$$P = \begin{bmatrix} 0.98 & 0.01 & 0.01 & 0 \\ 0 & 0.96 & 0.03 & 0.01 \\ 0 & 0 & 0.95 & 0.05 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Moreover, we let $c_R(3) = 40$ in Figure 24a and $c_R(3) = 5000$ in Figure 24b. The control-limit structure is violated in Figure 24a because, all else held equal, it is optimal to repair Asset 1 in conditions 1 and 3, but not in condition 2. In Figure 24b it is optimal to repair Asset 1 in conditions 1 and 2, but not in condition 3, i.e., it is optimal to abandon Asset 1 once it reaches condition 3.



(a) $c_R(3) = 40$ (b) $c_R(3) = 5000$

Figure 24: Excerpt of the optimal policies for the examples in Appendix B. The conditions of Theorem 10 are violated, and thus, the control-limit structure does not hold for Asset 1 in either Example (a) or (b). Moreover, in Example (b), assets are abandoned upon failure.

Appendix E Chapter 4.0: Downtime and Graph Centrality

In the examples of Section 4.5 all assets are in equally central locations, and thus, their average percentage of downtime is almost equal in the long-run. In this section, we study the relationship between the centrality of an asset's location and its downtime as well as how the downtime is affected by the graph structure.

Specifically, we consider six assets connected through four different graph configurations, namely, linear, grid, tree, and star configurations. We assume parameter values of $c_R(0) = 1$, $c_R(1) = 2$, $c_R(2) = 3$, $c_D = 1$, $c_T = 1$, $\lambda = 0.995$, and

$$P = \begin{bmatrix} 0.85 & 0.1 & 0.05 \\ 0 & 0.85 & 0.15 \\ 0 & 0 & 1 \end{bmatrix}.$$

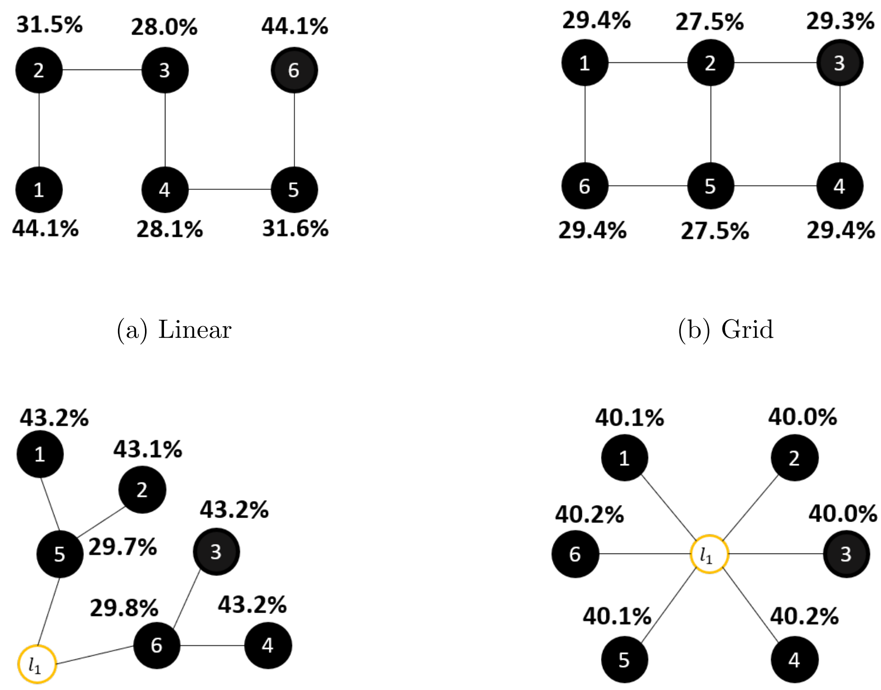
We run a simulation study for 600000 units of time and record the average percentage of downtime of all assets in each graph configuration; the results are depicted in Figure 25.

In Figure 25, first note that assets in less central locations incur more downtime compared to assets in more central locations. In fact, an asset's downtime decreases in its closeness centrality; recall equation (4.8). Second, assets on the grid graph have the least amount of downtime compared to assets on the other graph configurations, seemingly due to the high level of connectivity in the grid graph [144]. These results imply that assets' downtime is affected by both their relative centralities and network configuration.

Next, we present another example to examine the relationship between closeness centrality and downtime. Here, we let $c_R(0) = 1$, $c_R(1) = 2$, $c_R(2) = 3$, $c_D = 1$, $c_T = 1$, $= 0.995$, and

$$P = \begin{bmatrix} 0.95 & 0.03 & 0.02 \\ 0 & 0.95 & 0.05 \\ 0 & 0 & 1 \end{bmatrix}.$$

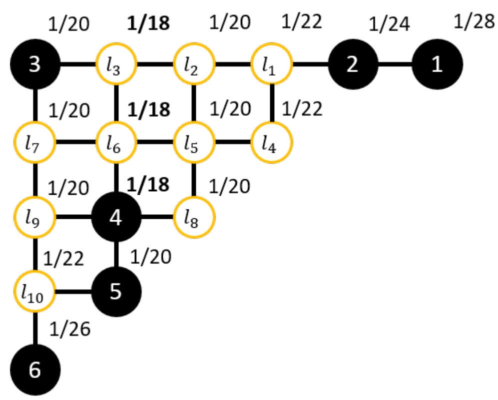
We run a simulation study for 2 million units of time to record the average percentage of downtime for each asset. Figure 26a depicts the graph configuration and Figure 26b plots the percentage of downtime against the closeness centrality of each asset.



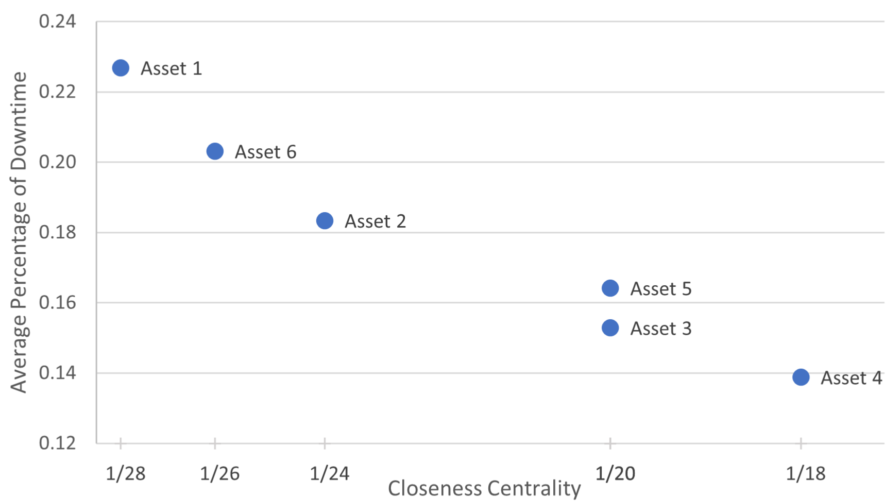
(a) Linear (b) Grid (c) Tree (d) Star

Figure 25: Average percentage of downtime for six assets in different graph configurations. The assets incur the lowest downtime in the grid graph due to its high connectivity. Moreover, in each graph, asset downtime is nondecreasing in closeness centrality.

The plot in Figure 26b implies that asset downtime decreases in the measure of closeness centrality. Furthermore, note that Asset 5 incurs a higher percentage of downtime compared to Asset 3 even though their measures of closeness centrality are equal. This difference can be explained by other measures of centrality such as eccentricity (i.e., the longest shortest path between the node of interest and all other asset nodes). In this example, Asset 5 has a larger eccentricity index compared to Asset 3.



(a) Graph for the example of Appendix C. Closeness centrality score, see equation (4.8), is labeled on the top of each node.



(b) Percentage of downtime versus closeness centrality of each asset

Figure 26: Plot of average percentage of downtime against the closeness centrality of six assets with the graph configuration depicted in (a). Asset downtime is decreasing in closeness centrality.

Appendix F Chapter 4.0: Deterioration Probability Matrices of Section 4.6

In the computational study of Section 4.6.3, we let the deterioration probability matrices be as follows:

$$P = \begin{bmatrix} 0.988 & 0.01 & 0.02 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.988 & 0.01 & 0.02 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.988 & 0.01 & 0.02 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.988 & 0.01 & 0.02 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.988 & 0.01 & 0.02 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.988 & 0.01 & 0.02 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.988 & 0.01 & 0.02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.988 & 0.01 & 0.02 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.988 & 0.12 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$P' = \begin{bmatrix} 0.98 & 0.012 & 0.005 & 0.002 & 0.001 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.98 & 0.012 & 0.005 & 0.002 & 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.98 & 0.012 & 0.005 & 0.002 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.98 & 0.012 & 0.005 & 0.002 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.98 & 0.012 & 0.005 & 0.002 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.98 & 0.012 & 0.005 & 0.002 & 0.001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.98 & 0.012 & 0.005 & 0.003 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.98 & 0.012 & 0.008 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.98 & 0.02 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$P'' = \begin{bmatrix} 0.97 & 0.021 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.97 & 0.021 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.97 & 0.021 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.97 & 0.021 & 0.006 & 0.002 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.97 & 0.021 & 0.006 & 0.002 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.97 & 0.021 & 0.006 & 0.002 & 0.001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.97 & 0.021 & 0.006 & 0.003 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.97 & 0.021 & 0.009 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.97 & 0.03 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$P''' = \begin{bmatrix} 0.95 & 0.041 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.95 & 0.041 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.95 & 0.041 & 0.006 & 0.002 & 0.001 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.95 & 0.041 & 0.006 & 0.002 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.95 & 0.041 & 0.006 & 0.002 & 0.001 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.95 & 0.041 & 0.006 & 0.002 & 0.001 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.95 & 0.041 & 0.006 & 0.003 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.95 & 0.041 & 0.009 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.95 & 0.05 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Bibliography

[1] D. T. Abdul-Malak and J. P. Kharoufeh. Optimally replacing multiple systems in a shared environment. *Probability in the Engineering and Informational Sciences*, 32(2):179, 2018.

[2] S. R. Agnihothri. A mean value analysis of the travelling repairman problem. *IIE Transactions*, 20(2):223–230, 1988.

[3] R. Ahmad and S. Kamaruddin. An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, 63(1):135–149, 2012.

[4] A. Ahrary, A. A. Nassiraei, and M. Ishikawa. A study of an autonomous mobile robot for a sewer inspection system. *Artificial Life and Robotics*, 11(1):23–27, 2007.

[5] R. Alanis, A. Ingolfsson, and B. Kolfal. A Markov chain model for an EMS system with repositioning. *Production and Operations Management*, 22(1):216–231, 2013.

[6] S. Alaswad and Y. Xiang. A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, 157:54–63, 2017.

[7] B. Angell. The Hidden Costs of Your Aging IT Infrastructure, 2017. Available from `https://blog.juriba.com/the-hidden-costs-of-an-aging-it-infrastructure`, last accessed on 5/2022.

[8] S. Arlinghaus. *Practical Handbook of Curve Fitting.* CRC Press, Boca Raton, Fl, 1994.

[9] M. J. Armstrong and D. A. Atkins. A note on joint optimization of maintenance and inventory. *IIE Transactions*, 30(2):143–149, 1998.

[10] T. Asfour, M. Waechter, L. Kaul, S. Rader, P. Weiner, S. Ottenhaus, R. Grimm, Y. Zhou, M. Grotz, and F. Paus. Armar-6: A high-performance humanoid for human-robot collaboration in real-world scenarios. *IEEE Robotics Automation Magazine*, 26(4):108–121, 2019.

[11] M. A. Bajestani and D. Banjevic. Calendar-based age replacement policy with dependent renewal cycles. *IIE Transactions*, 48(11):1016–1026, 2016.

[12] R. Barlow and L. Hunter. Optimum preventive maintenance policies. *Operations Research*, 8(1):90–100, 1960.

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

[14] P. R. Barnes, J. W. Van Dyke, B. W. McConnell, S. M. Cohn, and S. L. Purucker. The feasibility of replacing or upgrading utility distribution transformers during routine maintenance. Technical report, Oak Ridge National Lab., TN, 1995.

[15] Y. Barron and U. Yechiali. Generalized control-limit preventive repair policies for deteriorating cold and warm standby Markovian systems. *IISE Transactions*, 49(11):1031–1049, 2017.

[16] M. Ben-Daya, U. Kumar, and D. P. Murthy. *Introduction to maintenance engineering: modelling, optimization and management.* John Wiley & Sons, New York, NY, 2016.

[17] R. W. Bent and P. Van Hentenryck. Scenario-based planning for partially dynamic vehicle routing with stochastic customers. *Operations Research*, 52(6):977–987, 2004.

[18] O. Berman. Dynamic repositioning of indistinguishable service units on transportation networks. *Transportation Science*, 15(2):115–136, 1981.

[19] O. Berman. Repositioning of distinguishable urban service units on networks. *Computers & Operations Research*, 8(2):105–118, 1981.

[20] D. J. Bertsimas and G. Van Ryzin. A stochastic and dynamic vehicle routing problem in the Euclidean plane. *Operations Research*, 39(4):601–615, 1991.

[21] D. J. Bertsimas and G. Van Ryzin. Stochastic and dynamic vehicle routing in the Euclidean plane with multiple capacitated vehicles. *Operations Research*, 41(1):60–76, 1993.

[22] D. J. Bertsimas and G. Van Ryzin. Stochastic and dynamic vehicle routing with general demand and interarrival time distributions. *Advances in Applied Probability*, 25(4):947–978, 1993.

[23] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, NY, 2006.

[24] P. J. Boland. Periodic replacement when minimal repair costs vary with time. *Naval Research Logistics Quarterly*, 29(4):541–546, 1982.

[25] P. J. Boland and F. Proschan. Periodic replacement with increasing minimal repair costs at failure. *Operations Research*, 30(6):1183–1189, 1982.

[26] A. Boyle. How do you service a computer server? Send in the robots, Amazon patent says, 2018. Available from `https://www.geekwire.com/2018/service-computer-server-send-robots-amazon-patent-says/`, accessed on 4/2022.

[27] J. Brinkmann, M. W. Ulmer, and D. C. Mattfeld. Dynamic lookahead policies for stochastic-dynamic inventory routing in bike sharing systems. *Computers & Operations Research*, 106:260–279, 2019.

[28] J. Brinkmann, M. W. Ulmer, and D. C. Mattfeld. The multi-vehicle stochastic-dynamic inventory routing problem for bike sharing systems. *Business Research*, 13(1):69–92, 2020.

[29] E. Byon and Y. Ding. Season-dependent condition-based maintenance for a wind turbine using a partially observed markov decision process. *IEEE Transactions on Power Systems*, 25(4):1823–1834, 2010.

[30] F. Camci. The travelling maintainer problem: integration of condition-based maintenance with the travelling salesman problem. *Journal of the Operational Research Society*, 65(9):1423–1436, 2014.

[31] F. Camci. Maintenance scheduling of geographically distributed assets with prognostics information. *European Journal of Operational Research*, 245(2):506–516, 2015.

[32] A. M. Campbell and M. W. P. Savelsbergh. A decomposition approach for the inventory-routing problem. *Transportation Science*, 38(4):488–502, 2004.

[33] R. V. Canfield. Cost optimization of periodic preventive maintenance. *IEEE Transactions on Reliability*, 35(1):78–81, April 1986.

[34] Casey, B. Advice For Maintaining Hydraulic Accumulators, 2009. Available from `https://www.efficientplantmag.com/2009/11/measuring-maintenance-effectiveness-part-iii-497/`, last accessed on 5/2022.

[35] C. R. Cassady, I. M. Iyoob, K. Schneider, and E. A. Pohl. A generic model of equipment availability under imperfect maintenance. *IEEE Transactions on Reliability*, 54(4):564–571, 2005.

[36] M. Chen, S. Nakamura, and T. Nakagawa. Replacement and preventive maintenance models with random working times. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 93(2):500–507, 2010.

[37] N. Chen, Z. S. 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.

[38] Y. Chen, P. Cowling, F. Polack, S. Remde, and P. Mourdjis. Dynamic optimisation of preventative and corrective maintenance schedules for a large scale urban drainage system. *European Journal of Operational Research*, 257(2):494–510, 2017.

[39] Y. H. Chien. The effect of a pro-rata rebate warranty on the age replacement policy with salvage value consideration. *IEEE Transactions on Reliability*, 59(2):383–392, 2010.

[40] S. A. Chowdhery and M. Bertoni. Modeling resale value of road compaction equipment: a data mining approach. *IFAC-PapersOnLine*, 51(11):1101–1106, 2018.

[41] B. Christiansen. 5 Important Maintenance Metrics and How to Use Them, 2018. Available from `https://www.maintworld.com/Applications/5-Important-Maintenance-Metrics-and-How-To-Use-Them`, last accessed on 5/2022.

[42] L. C. Coelho, J.-F. Cordeau, and G. Laporte. Thirty years of inventory routing. *Transportation Science*, 48(1):1–19, 2014.

[43] A. Crespo Márquez, P. Moreu de León, J. F. Gómez Fernández, C. Parra Márquez, and M. López Campos. The maintenance management framework: A practical view to maintenance management. *Journal of Quality in Maintenance Engineering*, 15(2):167–178, 2009.

[44] Data Insights Partner. Inspection and Maintenance Robots Market : Global Industry Analysis, Size, Share, Growth, Trends and Forecast, 2020-2028, 2021. Available from `https://datainsightspartner.com/report/inspection-and-maintenance-robots-market/963`, accessed on 2/2022.

[45] B. de Jonge and P. A. Scarf. A review on maintenance optimization. *European Journal of Operational Research*, 285(3):805–824, 2020.

[46] B. de Jonge, R. Teunter, and T. Tinga. The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance. *Reliability Engineering & System Safety*, 158:21–30, 2017.

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

[48] F. Ding and Z. Tian. Opportunistic maintenance for wind farms considering multi-level imperfect maintenance thresholds. *Renewable Energy*, 45:175–182, 2012.

[49] C. Drent, M. O. Keizer, and G.-J. van Houtum. Dynamic dispatching and repositioning policies for fast-response service networks. *European Journal of Operational Research*, 285(2):583–598, 2020.

[50] S. O. Duffuaa, A. Raouf, and J. D. Campbell. *Planning and Control of Maintenance Systems: Modeling and Analysis*. John Wiley & Sons, New York, NY, 1999.

[51] D. Fan, Y. Ren, Q. Feng, B. Zhu, Y. Liu, and Z. Wang. A hybrid heuristic optimization of maintenance routing and scheduling for offshore wind farms. *Journal of Loss Prevention in the Process Industries*, 62:103949, 2019.

[52] H. Fan, S. AbouRizk, H. Kim, and O. Zaïane. Assessing residual value of heavy construction equipment using predictive data mining model. *Journal of Computing in Civil Engineering*, 22(3):181–191, 2008.

[53] W. D. Fan, M. D. Gemar, and R. Machemehl. Equipment replacement decision making: Opportunities and challenges. *Journal of the Transportation Research Forum*, 52(3), 2013.

[54] L. Fangxing and R. E. Brown. A cost-effective approach of prioritizing distribution maintenance based on system reliability. *IEEE Transactions on Power Delivery*, 19(1):439–441, 2004.

[55] T. A. Feo and J. F. Bard. Flight scheduling and maintenance base planning. *Management Science*, 35(12):1415–1432, 1989.

[56] M. Finkelstein and S. Eryilmaz. On optimal maintenance of degrading multistate systems with state-dependent cost of repair. *Applied Stochastic Models in Business and Industry*, 37(4):790–801, 2021.

[57] M. Finkelstein, M. Shafiee, and A. N. Kotchap. Classical optimal replacement strategies revisited. *IEEE Transactions on Reliability*, 65(2):540–546, 2016.

[58] D. M. Frangopol and M. Liu. Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost. *Structure and Infrastructure Engineering*, 3(1):29–41, 2007.

[59] G. Gallucci, W. Swartz, and F. Hackerman. Brief reports: Impact of the wait for an initial appointment on the rate of kept appointments at a mental health center. *Psychiatric Services*, 56(3):344–346, 2005.

[60] A. Goel and F. Meisel. Workforce routing and scheduling for electricity network maintenance with downtime minimization. *European Journal of Operational Research*, 231(1):210–228, 2013.

[61] H. Golmakani and M. Pouresmaeeli. Optimal replacement threshold and inspection interval for condition-based maintenance with variable failure cost. In *2012 IEEE International Conference on Industrial Engineering and Engineering Management*, pages 1944–1948. IEEE, 2012.

[62] M. J. Havinga and B. de Jonge. Condition-based maintenance in the cyclic patrolling repairman problem. *International Journal of Production Economics*, 222:107497, 2020.

[63] K. He, L. M. Maillart, and O. A. Prokopyev. Scheduling preventive maintenance as a function of an imperfect inspection interval. *IEEE Transactions on reliability*, 64(3):983–997, 2015.

[64] K. He, L. M. Maillart, and O. A. Prokopyev. Optimal planning of unpunctual preventive maintenance. *IISE Transactions*, 49(2):127–143, 2017.

[65] K. He, L. M. Maillart, and O. A. Prokopyev. Optimal sequencing of heterogeneous, non-instantaneous interventions. *Annals of Operations Research*, 276(1):109–135, 2019.

[66] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds. Wireless sensor networks for condition monitoring in the railway industry: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 16(3):1088–1106, 2015.

[67] M. Hoffman, E. Song, M. P. Brundage, and S. Kumara. Online improvement of condition-based maintenance policy via monte carlo tree search. *IEEE Transactions on Automation Science and Engineering*, pages 1–12, 2021.

[68] C. C. Huang, W. T. Ziemba, and A. Ben-Tal. Bounds on the expectation of a convex function of a random variable: With applications to stochastic programming. *Operations Research*, 25(2):315–325, 1977.

[69] C. A. Irawan, D. Ouelhadj, D. Jones, M. Stålhane, and I. B. Sperstad. Optimisation of maintenance routing and scheduling for offshore wind farms. *European Journal of Operational Research*, 256(1):76–89, 2017.

[70] M. Jamil, R. Batta, and D. M. Malon. The traveling repairperson home base location problem. *Transportation Science*, 28(2):150–161, 1994.

[71] A. K. Jardine, D. Lin, and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7):1483–1510, 2006.

[72] C. Jia and C. Zhang. Joint optimization of maintenance planning and workforce routing for a geographically distributed networked infrastructure. *IISE Transactions*, 52(7):732–750, 2020.

[73] S. K. Keneally, M. J. Robbins, and B. J. Lunday. A Markov decision process model for the optimal dispatch of military medical evacuation assets. *Health Care Management Science*, 19(2):111–129, 2016.

[74] R. Khan, A. B. Mad, K. Osman, and M. A. A. Aziz. Maintenance management of aging oil and gas facilities. In *Maintenance Management*, chapter 5. IntechOpen, London, 2019.

[75] A. J. Kleywegt, V. S. Nori, and M. W. P. Savelsbergh. The stochastic inventory routing problem with direct deliveries. *Transportation Science*, 36(1):94–118, 2002.

[76] A. J. Kleywegt, V. S. Nori, and M. W. P. Savelsbergh. Dynamic programming approximations for a stochastic inventory routing problem. *Transportation Science*, 38(1):42–70, 2004.

[77] C. H. Kriebel, A. A. Atkinson, and H. W. H. Zia. Optimal investment, pricing and replacement of computer resources. *Naval Research Logistics Quarterly*, 24(4):537–547, 1977.

[78] M. Kurt and J. P. Kharoufeh. Monotone optimal replacement policies for a Markovian deteriorating system in a controllable environment. *Operations Research Letters*, 38(4):273–279, 2010.

[79] A. Larsen, O. Madsen, and M. M. Solomon. Partially dynamic vehicle routing-models and algorithms. *Journal of the Operational Research Society*, 53(6):637–646, 2002.

[80] A. Larsen, O. B. Madsen, and M. M. Solomon. The a priori dynamic traveling salesman problem with time windows. *Transportation Science*, 38(4):459–472, 2004.

[81] D. J. Leatham and T. G. Baker. Empirical estimates of the effects of inflation on salvage values, cost and optimal replacement of tractors and combines. *North Central Journal of Agricultural Economics*, 3(2):109–117, 1981.

[82] V. J. Lee, A. Earnest, M. I. Chen, and B. Krishnan. Predictors of failed attendances in a multi-specialty outpatient centre using electronic databases. *BMC Health Services Research*, 5(1):51, 2005.

[83] P. Liljebäck and R. Mills. Eelume: A flexible and subsea resident IMR vehicle. In *Oceans 2017-Aberdeen*, pages 1–4. IEEE, 2017.

[84] J. Lister. Reasons to Replace Business Computers, 2013. Available from `https://smallbusiness.chron.com/reasons-replace-business-computers-75262.html`, last accessed on 5/2022.

[85] E. López-Santana, R. Akhavan-Tabatabaei, L. Dieulle, N. Labadie, and A. L. Medaglia. On the combined maintenance and routing optimization problem. *Reliability Engineering & System Safety*, 145:199–214, 2016.

[86] G. Lucko. Modeling the residual market value of construction equipment under changed economic conditions. *Journal of Construction Engineering and Management*, 137(10):806–816, 2011.

[87] L. M. Maillart and X. Fang. Optimal maintenance policies for serial, multi-machine systems with non-instantaneous repairs. *Naval Research Logistics*, 53(8):804–813, 2006.

[88] MarketsandMarkets. Machine Condition Monitoring Market, Report Code: SE 2739, 2021. Available from `https://www.marketsandmarkets.com/Market-Reports/machine-health-monitoring-market-29627363.html`, accessed on 4/2022.

[89] M. Marom. What to do with aging medium voltage switchgear?, 2019. Available from `https://blog.se.com/services/2019/04/26/what-to-do-with-aging-medium-voltage-switchgear/`, last accessed on 5/2022.

[90] M. S. Maxwell, E. C. Ni, C. Tong, S. G. Henderson, H. Topaloglu, and S. R. Hunter. A bound on the performance of an optimal ambulance redeployment policy. *Operations Research*, 62(5):1014–1027, 2014.

[91] M. S. Maxwell, M. Restrepo, S. G. Henderson, and H. Topaloglu. Approximate dynamic programming for ambulance redeployment. *INFORMS Journal on Computing*, 22(2):266–281, 2010.

[92] A. Mérigaud and J. V. Ringwood. Condition-based maintenance methods for marine renewable energy. *Renewable and Sustainable Energy Reviews*, 66:53–78, 2016.

[93] C. Morris. Will The Great Resignation Lead to a Rise of the Robots?, 2021. Available from `https://www.nasdaq.com/articles/will-the-great-resignation-lead-to-a-rise-of-the-robots`, accessed on 2/2022.

[94] C. Mueller. Market for Professional and Domestic Service Robots Booms in 2018, 2019. Available from `https://ifr.org/post/market-for-professional-and-domestic-service-robots-booms-in-2018`, accessed on 4/2022.

[95] T. Nakagawa. Sequential imperfect preventive maintenance policies. *IEEE Transactions on Reliability*, 37(3):295–298, Aug 1988.

[96] T. Nakagawa. *Maintenance Theory of Reliability*. Springer Science & Business Media, London, UK, 2006.

[97] T. Nakagawa, X. Zhao, and W. Y. Yun. Optimal age replacement and inspection policies with random failure and replacement times. *International Journal of Reliability, Quality and Safety Engineering*, 18(05):405–416, 2011.

[98] M. Newman. *Networks*. Oxford University Press, Oxford, UK, 2018.

[99] H. S. H. Nguyen, P. Do, H. Vu, and B. Iung. Dynamic maintenance grouping and routing for geographically dispersed production systems. *Reliability Engineering & System Safety*, 185:392–404, 2019.

[100] N. Osadchiy and D. KC. Are patients patient? The role of time to appointment in patient flow. *Production and Operations Management*, 26(3):469–490, 2017.

[101] J. Pieters. Overdue Maintenance Could Lead to Traffic Disaster: Study, 2019. Available from `https://nltimes.nl/2019/06/03/overdue-maintenance-lead-traffic-disaster-study`, last accessed on 5/2022.

[102] L. M. Pintelon and L. F. Gelders. Maintenance management decision making. *European Journal of Operational Research*, 58(3):301–317, 1992.

[103] S. S. Ponnaluru, T. L. Marsh, and M. Brady. Spatial price analysis of used construction equipment: the case of excavators. *Construction Management and Economics*, 30(11):981–994, 2012.

[104] W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, Newyork, NY, 2007.

[105] A. Prajapati, J. Bechtel, and S. Ganesan. Condition based maintenance: a survey. *Journal of Quality in Maintenance Engineering*, 18(4):384–400, 2012.

[106] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Newyork, NY, 2014.

[107] M. Rashidnejad, S. Ebrahimnejad, and J. Safari. A bi-objective model of preventive maintenance planning in distributed systems considering vehicle routing problem. *Computers & Industrial Engineering*, 120:360–381, 2018.

[108] H. Remer. Why You Need a Computer Replacement Schedule, 2015. Available from https://www.businessmagazinegainesville.com/why-you-need-a-computer-replacement-schedule/, last accessed on 5/2022.

[109] Remi. Aging equipment maintenance costs - part 2, 2019. Available from https://theremigroup.com/aging-equipment-2/, last accessed on 5/2022.

[110] G. Roesler, P. Jaffe, and G. Henshaw. Inside DARPA's mission to send a repair robot to geosynchronous orbit, 2017. Available from https://spectrum.ieee.org/inside-darpas-mission-to-send-a-repair-robot-to-geosynchronous-orbit, accessed on 3/2022.

[111] S. M. Ross. *Applied Probability Models with Optimization Applications*. Dover Books on Advanced Mathematics. Dover Publications, Mineola, NY, 1970.

[112] N. Safaei and A. K. Jardine. Aircraft routing with generalized maintenance constraints. *Omega*, 80:111–122, 2018.

[113] A. Salonen and M. Deleryd. Cost of poor maintenance: A concept for maintenance performance improvement. *Journal of Quality in Maintenance Engineering*, 17(1):63–73, 2011.

[114] G. H. Sandler. *System Reliability Engineering*. Prentice-Hall Englewood, Cliffs, NJ, 1963.

[115] P. A. Scarf. A framework for condition monitoring and condition based maintenance. *Quality Technology & Quantitative Management*, 4(2):301–312, 2007.

[116] R. L. Scheaffer. Optimum age replacement policies with an increasing cost factor. *Technometrics*, 13(1):139–144, 1971.

[117] A. H. Schrotenboer, M. A. uit het Broek, B. Jargalsaikhan, and K. J. Roodbergen. Coordinating technician allocation and maintenance routing for offshore wind farms. *Computers & Operations Research*, 98:185–197, 2018.

[118] Y. Segawa, M. Ohnishi, and T. Ibaraki. Optimal minimal-repair and replacement problem with age dependent cost structure. *Computers and Mathematics with Applications*, 24(1):91–101, 1992.

[119] G. Sethuraman and J. Jesintha. All banana trees are graceful. *Advances in Applied Discrete Mathematics*, 4:53–64, 2009.

[120] S. H. Sheu and W. S. Griffith. Optimal age-replacement policy with age-dependent minimal-repair and random-leadtime. *IEEE Transactions on Reliability*, 50(3):302–309, 2001.

[121] Y. Shi, Y. Xiang, and M. Li. Optimal maintenance policies for multi-level preventive maintenance with complex effects. *IISE Transactions*, 51(9):999–1011, 2019.

[122] D. Simchi-Levi, X. Chen, and J. Bramel. *The logic of logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*. Springer Series in Operations Research. Springer, New York, NY, 2005.

[123] K. T. Talluri. The four-day aircraft maintenance routing problem. *Transportation Science*, 32(1):43–53, 1998.

[124] S. Telford, M. I. Mazhar, and I. Howard. Condition based maintenance (CBM) in the oil and gas industry: an overview of methods and techniques. In *International Conference on Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia Jan*, pages 22–24, 2011.

[125] B. W. Thomas. Waiting strategies for anticipating service requests from known customer locations. *Transportation Science*, 41(3):319–331, 2007.

[126] Z. Tian, T. Jin, B. Wu, and F. Ding. Condition based maintenance optimization for wind power generation systems under continuous monitoring. *Renewable Energy*, 36(5):1502–1509, 2011.

[127] C. Tilquin and R. Cléroux. Periodic replacement with minimal repair at failure and general cost function. *Journal of Statistical Computation and Simulation*, 4(1):63–77, 1975.

[128] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer Science & Business Media, London, UK, 2013.

[129] A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao. Finding groups with maximum betweenness centrality. *Optimization Methods and Software*, 32(2):369–399, 2017.

[130] A. Veremyev, O. A. Prokopyev, and E. L. Pasiliao. Finding critical links for closeness centrality. *INFORMS Journal on Computing*, 31(2):367–389, 2019.

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

[132] X. Wang, L. Li, and M. Xie. An unpunctual preventive maintenance policy under two-dimensional warranty. *European Journal of Operational Research*, 282(1):304–318, 2020.

[133] J. Whittle, G. Schectman, N. Lu, B. Baar, and M. F. Mayo-Smith. Relationship of scheduling interval to missed and cancelled clinic appointments. *Journal of Ambulatory Care Management*, 31(04):290 – 302, 2008.

[134] T. Xia, Y. Dong, E. Pan, M. Zheng, H. Wang, and L. Xi. Fleet-level opportunistic maintenance for large-scale wind farms integrating real-time prognostic updating. *Renewable Energy*, 163:1444–1454, 2021.

[135] P. Xu, R. Liu, Q. Sun, and L. Jiang. Dynamic-time-warping-based measurement data alignment model for condition-based railroad track maintenance. *IEEE Transactions on Intelligent Transportation Systems*, 16(2):799–812, 2014.

[136] R. H. Yeh, G. C. Chen, and M. Y. Chen. Optimal age-replacement policy for nonrepairable products under renewing free-replacement warranty. *IEEE Transactions on Reliability*, 54(1):92–97, 2005.

[137] X. Zhao, K. N. Al-Khalifa, A. M. Hamouda, and T. Nakagawa. First and last triggering event approaches for replacement with minimal repairs. *IEEE Transactions on Reliability*, 65(1):197–207, March 2016.

[138] X. Zhao, K. N. Al-Khalifa, A. M. Hamouda, and T. Nakagawa. Age replacement models: A summary with new perspectives and methods. *Reliability Engineering & System Safety*, 161:95–105, 2017.

[139] X. Zhao and T. Nakagawa. Optimization problems of replacement first or last in reliability theory. *European Journal of Operational Research*, 223(1):141–149, 2012.

[140] X. Zhao and T. Nakagawa. *Age Replacement Models with Random Works*, pages 227–248. Springer, London, UK, 2016.

[141] X. Zhao, C. Qian, and T. Nakagawa. Comparisons of replacement policies with periodic times and repair numbers. *Reliability Engineering & System Safety*, 168:161–170, 2017.

[142] Z. Zheng, M. Li, X. Xiao, and J. Wang. Coordinated resource provisioning and maintenance scheduling in cloud data centers. In *2013 Proceedings IEEE INFOCOM*, pages 345–349, 2013.

[143] X. Zhou and M. Yu. Semi-dynamic maintenance scheduling for multi-station series systems in multi-specification and small-batch production. *Reliability Engineering & System Safety*, 195:106753, 2020.

[144] R. Żochowska and P. Soczówka. Analysis of selected transportation network structures based on graph measures. *Zeszyty Naukowe. Transport/Politechnika Śląska*, 2018.