Condition-Based Production for Stochastically Deteriorating Systems: Optimal Policies and Learning

Collin Drent, Melvin Drent

School of Industrial Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands,
{c.drent, m.drent}@tue.nl

Joachim Arts

Luxembourg Centre for Logistics and Supply Chain Management, University of Luxembourg, Luxembourg-City, Luxembourg,
joachim.arts@uni.lu

Problem Definition: Production systems deteriorate stochastically due to usage and may eventually break down, resulting in high maintenance costs at scheduled maintenance moments. This deterioration behavior is affected by the system's production rate. While producing at a higher rate generates more revenue, the system may also deteriorate faster. Production should thus be controlled dynamically to trade-off deterioration and revenue accumulation in between maintenance moments. We study systems for which the relation between production and deterioration is known and the same for each system as well as systems for which this relation differs from system to system and needs to be learned on-the-fly. The decision problem is to find the optimal production policy given planned maintenance moments (operational) and the optimal interval length between such maintenance moments (tactical). **Methodology/Results:** For systems with a known production-deterioration relation, we cast the operational decision problem as a continuous-time Markov decision process and prove that the optimal policy has intuitive monotonic properties. We also present sufficient conditions for the optimality of bang-bang policies and we partially characterize the structure of the optimal interval length, thereby enabling efficient joint optimization of the operational and tactical decision problem. For systems that exhibit variability in their production-deterioration relations, we propose a Bayesian procedure to learn the unknown deterioration rate under any production policy. Numerical studies indicate that on average across a wide range of settings (i) condition-based production increases profits by 50% compared to static production, (ii) integrating condition-based production and maintenance decisions increases profits by 21% compared to the state-of-the-art sequential approach, and (iii) our Bayesian approach performs close – especially in the bang-bang regime – to an Oracle policy that knows each system's production-deterioration relation. **Managerial Implications:** Production should be adjusted dynamically based on real-time condition monitoring and the tactical maintenance planning should anticipate and integrate these operational decisions. Our proposed framework assists managers to do so optimally.

Key words: optimal production control, maintenance, Markov decision process, Bayesian learning, data-driven operations

1. Introduction

The modern manufacturing industry relies heavily and increasingly on advanced technical systems in their daily operations. These systems – for instance lithography systems in the semiconductor industry, wind turbines in wind farms, and assembly equipment in production lines – deteriorate due to usage and may break down as a result of such deterioration, leading to extremely high costs in the event of prolonged downtime. For example, unplanned downtime of lithography equipment

can cost up to 72,000 euros per hour (Lamghari-Idrissi 2021). The global manufacturing industry, which includes the semiconductor industry, faces an estimated annual cost of around $50 billion due to unplanned system downtimes and their consequences (Coleman et al. 2017).

Motivated by such cost figures, there has been a considerable amount of research on maintenance optimization, where the main goal is to determine when to preventively maintain a system based on its condition in order to prevent costly failures and their associated downtimes (see Wang (2002) and De Jonge and Scarf (2020) for excellent overviews of the field). A common assumption in this field is that the underlying deterioration of a system is exogenous; that is, the deterioration is not influenced by factors other than by its own randomness. However, for many advanced technical systems in the manufacturing industry, the deterioration behavior of its components is often affected by the production load applied to the system itself. For example, gearboxes and generators in a wind turbine deteriorate faster under higher rotational speeds (Feng et al. 2013, Macquart and Maheri 2019) which can be adjusted instantly and remotely by controlling the wind turbine's pitch (Spinato et al. 2009, Zhang et al. 2015). Other examples include cutting tools in high-speed cutting machinery that wear faster under higher speeds (Dolinšek et al. 2001), conveyor belts in production lines that are increasingly likely to fail as their rotational speeds are increased (Nourelfath and Yalaoui 2012), stamping machines that deteriorate faster as the stamping speed increases (Hao et al. 2015), and rapid gravity filters in waterworks that deteriorate faster as the filtration rate is increased (Sun et al. 2019). There is a caveat to this: Production loads that lead to less deterioration accumulation typically also come with less revenue accumulation. Indeed, wind turbines generate less wind energy when their rotational speeds are lower. Production lines and cutting tools have less output when conveyor belts and cutting speeds are slowed down, respectively. Production decisions in make-to-stock environments are not driven by order deadlines but by revenue targets and equipment maintenance requirements. It is in these environments that our contribution will apply best.

Systems usually fail when components deteriorate beyond a critical level. As such, failures can be prevented before they occur by making instant maintenance decisions based on the component's deterioration level. (Hence most research has focused on the decision when to preventively maintain a system based on its condition.) However, in practice, maintenance planning has limited flexibility and cannot be done last minute because arranging the logistic support takes time, particularly for remotely located systems (e.g., wind turbines in offshore wind farms). Moreover, it is inconvenient and costly to shut down an entire system each time an individual component needs maintenance. Maintenance intervals across a set of assets/components are therefore often fixed to reduce the fixed cost of shutting down an asset and bringing all the maintenance resources together. The classical example of this is the block maintenance policy studied by Barlow and Hunter (1960), forming the

basis for more recent work, such as that of Van Dijkhuizen and van Harten (1997), Arts and Basten (2018), and Uit het Broek et al. (2020). For these reasons, in practice, maintenance activities are usually performed at planned maintenance moments that are scheduled periodically and well in advance. At such planned maintenance moments, the entire system is deliberately shut down and components that require maintenance are maintained.

When these planned maintenance moments are scheduled in advance, another strategy that decision makers (DMs) can employ to reduce maintenance costs and increase revenues, is to control the deterioration behavior in between such maintenance moments. Specifically, slowing down the deterioration rate by decreasing the production rate when a component's deterioration is close to its critical level leads to fewer failures between consecutive maintenance moments. On the other hand, if components deteriorate slowly, DMs can increase the production rate, thereby accumulating a higher revenue without running the risk of system failure before the next maintenance moment.

This operational strategy has become a viable option due to the increasing and improved availability of real-time data stemming from ubiquitous sensors installed in modern equipment, integrated in the so-called Internet-of-Things (Coopers 2014, Blackwell et al. 2017). It particularly enables (i) remote accrual of deterioration data, and (ii) continuous remote control of production operations at near-zero costs and in real time (Olsen and Tomlin 2020). This novel idea was recently proposed by Uit het Broek et al. (2020), who coined the term condition-based production to describe this strategy. The authors consider condition-based production rate decisions for systems with an adjustable production rate that directly affects the deterioration rate. There is an inherent trade-off to this concept: A lower production rate has both a lower deterioration and revenue rate, while a higher production rate has both a higher deterioration and revenue rate. By formulating a continuous-time optimal control problem, Uit het Broek et al. (2020) are able to study this trade-off and establish structural insights for the case that deterioration is deterministic. As yet, there are no analytical results for stochastic deterioration processes despite significant practical relevance.

In this paper, we study such optimal condition-based production policies. Contrary to the existing literature, we consider stochastically deteriorating systems, and our objective is to characterize the structure of the optimal policy. To this end, we model the stochastic deterioration as a Poisson process whose uncontrolled intensity is equal to a certain base rate. Modeling deterioration as a Poisson process is a common and empirically validated approach in the literature for systems subject to so-called stochastic shock deterioration, where systems only deteriorate due to shocks that occur at random points in time (Esary and Marshall 1973, Sobczyk 1987). The base rate represents the rate at which the system deteriorates under normal, uncontrolled operating conditions. A DM can continuously adjust production which leads to instantaneous changes in the revenue rate and

the rate of the Poisson process (proportional to the base rate). We shall discuss this in more detail in §3, but conceptually, we have that a higher (lower) production rate leads to higher (lower) revenue and deterioration rates. By employing such a condition-based production policy, the DM can steer the deterioration as she nears the planned maintenance moment, at which a maintenance cost depending on the deterioration state is incurred.

We first consider the case where each newly installed system has the same fixed base rate that is a-priori known to the DM. We model this problem as a continuous-time Markov decision process (MDP) when the planned maintenance interval is fixed and rigorously analyze its Hamilton-Jacobi-Bellman (HJB) equations to characterize the monotonic behavior of the optimal policy. We also provide easily verifiable conditions that are sufficient for a so-called bang-bang policy (production is either set to the maximum or off) to be optimal, thereby reducing the action space of the continuous-time MDP from a compact set with infinitely many actions to a discrete set with only two actions. We then study the case where each newly installed system has a different base rate that is unknown a-priori and must be learned in real time, that is, the population of systems exhibits unit-to-unit variability in their base rates. For this case, we develop an easy-to-implement Bayesian heuristic, that mimics the structure of the optimal policy for the known base rate case. In this scenario, a DM needs to learn the a-priori unknown base rate based on the chosen production policy and the deterioration she observes as a consequence of that policy.

Finally, we explore the tactical trade-off due to maintenance costs levied at a planned maintenance moment. Specifically, scheduling planned maintenance moments too frequently leads to unnecessary maintenance activities and associated costs, while a very long maintenance interval both increases the risk of a failed system and decreases the revenue accumulation as the DM is then forced to decrease the production rate. This implies that in deciding upon the length of the interval between planned maintenance moments (a tactical decision), the optimal condition-based production policy (operational decisions) needs to be taken into account. We study this trade-off and show that the optimization of the maintenance interval is a unimodal optimization problem.

The main contributions of this paper are fourfold:

1. We are the first to analytically establish the structure of the optimal condition-based production policy under stochastic deterioration. The structure of the optimal policy has intuitive appeal: Decrease production if the system's condition is close to failure, and increase production if the planned maintenance moment is nearing or when the base rate is lower. We also establish sufficient conditions for when the optimal policy has a bang-bang structure. As a byproduct of our analysis, we obtain new submodularity preservation properties that may have merit themselves in future work in the broad field of monotonic optimal control theory.

2. We show that under the optimal condition-based production policy, the length of the interval between planned maintenance moments can be easily optimized. This allows for the optimal integration of condition-based production on an operational level and the length of planned maintenance intervals on a tactical level.

3. We complement our theoretical results with an extensive numerical study, which demonstrates that on average across a wide range of practical settings: (i) condition-based production policies lead to 50% profit increases compared to optimal static policies, and (ii) integrating maintenance decisions and condition-based production to 21% profit increases compared to the state-of-the-art sequential approach. Further comparative static analyses offer insights into the driving factors behind these profit increases, and consequently highlight practical settings that are associated with higher or lower increases.

4. For systems that exhibit unit-to-unit variability in their base rates, we propose a Bayesian procedure to tractably learn this unknown rate under any production policy, and we use it to build a heuristic policy. In an extensive simulation study, we show that this heuristic performs close to an Oracle policy that knows the base rate a-priori. Notably, our results highlight a key insight: The heuristic yields particularly good performance in the bang-bang regime, and this performance remains robust against unit-to-unit variability.

The rest of the paper is structured as follows. In §2, we review related literature. We present the problem formulation in §3. In §4, we present the main analytical results for populations of systems whose base rates are fixed and known. §5 presents a numerical study that underscores the practical value of our theoretical results. In §6, we study system populations with unit-to-unit variability in their base rates. Finally, §7 provides concluding remarks. All proofs are relegated to Appendix A.

2. Literature Review

Studies on production control under uncertainty as well as on maintenance optimization are abundant in literature. Sethi et al. (2002) and Mula et al. (2006) provide excellent overviews of the former field while Alaswad and Xiang (2017) and De Jonge and Scarf (2020) do so for the latter. The intersection of these two fields has received considerably less attention in the literature. In what follows we discuss studies on this intersection that are most relevant to our work. Additionally, since our methodology exhibits similarities with modeling approaches in revenue management, we conclude by discussing and contrasting our work with this literature stream too.

Most of the early works focus on production dependent failures. In such settings, production equipment resides in either a failed or non-failed state, and the probability of failing depends on the production rate. Liberopoulos and Caramanis (1994) were among the first to study such a setting. They consider a make-to-stock production system where the system's lifetime has an exponential

distribution whose rate depends on the production rate. Upon failure, corrective maintenance is needed, which also takes an exponentially distributed amount of time. Liberopoulos and Caramanis (1994) focus on the cost performance evaluation of so-called threshold polices. Under such a policy, the production rate is governed by thresholds on the inventory level, where higher inventory levels prescribe lower production rates. Many follow-up studies that depart from this work have appeared since then. While some authors have focused on providing conditions under which certain threshold policies are optimal, typically dependent on the production-failure relationship (e.g., Hu and Xiang 1995, Martinelli 2010), others have extended the model by incorporating the option to perform preventive maintenance that is cheaper than corrective maintenance (e.g., Boukas et al. 1995, Boukas and Liu 2001), by considering multiple production machines (e.g., Francie et al. 2014), or by letting the failure rate depend on other variables, for instance the equipment's age (e.g., Zied et al. 2011). Although there are practical settings where production equipment suddenly ceases to operate from a normal operating state, most production equipment deteriorates gradually during production. For that reason, research has focused on extending the research line of production dependent failures to models in which the deterioration of the equipment can be described by more than one non-failed state, leading to so-called production dependent deterioration.

Early works on production dependent deterioration can be found in Conrad and McClamroch (1987) and Rishel (1989). Conrad and McClamroch (1987) study a drilling machine with controllable speed whose condition deteriorates according to a speed-dependent Brownian motion. Rishel (1989) studies a similar problem, but considers a wear process that is monotone. Similar to our paper, both works let the deterioration and revenue rate depend on the control variable, but unlike our work, they incorporate a deterioration-dependent cost rate and wish to maximize the expected profit over the system's random lifespan. These works assume that maintenance can be performed instantaneously at any time. This assumption is often not tenable in practice since maintenance – especially for remotely located systems – is usually planned well in advance and therefore permits limited flexibility, if any. In our work, we take this into account and consider the case where maintenance can only be performed at pre-specified planned times; we then wish to maximize the expected profit over such a planning horizon. Moreover, Conrad and McClamroch (1987) and Rishel (1989) do not provide analytical results regarding the structure of the optimal policy. Since these early works, researchers have extended production dependent deterioration models with the inclusion of maintenance activities to improve the equipment's condition (e.g., Sloan and Shanthikumar 2000, 2002, Kazaz and Sloan 2008, Batun and Maillart 2012, Kazaz and Sloan 2013, Uit het Broek et al. 2021), but again with few explicit structural results.

The work on production dependent deterioration that is most relevant to ours can be found in Uit het Broek et al. (2020). Like us, they study how to dynamically adjust the production rate

over a finite planning horizon that ends with a maintenance moment, where a higher production rate leads to more revenue, but also to a faster deterioration rate, potentially leading to higher maintenance costs at the end of the planning horizon. The authors formulate an optimal control problem and characterize the structure of the optimal condition-based production policy under the assumption that deterioration is deterministic and its parameters fully known. When deterioration is stochastic, the authors numerically explore the optimal policy but do not report on structural results. Our work extends the work of Uit het Broek et al. (2020) in three important ways. Firstly, we establish the structural properties of the optimal condition-based production policy for stochastically deteriorating systems; these properties are found to differ significantly from those under deterministic deterioration. Secondly, we study the case in which each system's deterioration process is characterized by an a-priori unknown parameter that must be learned on-the-fly, which gears our model even more towards practice. Indeed, recent practice-based research has shown that deterioration processes of the type studied in this work may exhibit unit-to-unit variability (Drent et al. 2023a), meaning that the parameters of the deterioration process of each newly installed system differ from previous ones. Finally, we show that the interval length between planned maintenance moments can be easily optimized under the optimal condition-based production policy. This allows for the optimal integration of condition-based production decisions on an operational level and maintenance decisions on a tactical level, which has not been studied before.

Methodologically, our modeling approach shares some similarities with approaches to model the canonical dynamic pricing problem in revenue management, where a DM seeks to maximize revenue by pricing a finite number of products within a fixed time frame. In this problem, introduced by Gallego and Van Ryzin (1994), customers arrive according to a non-homogeneous Poisson process whose intensity depends on the chosen price. They formulate this problem as an intensity control problem, also leading to a set of HJB equations. Their work instigated a large stream of literature in which various extensions are studied, such as multiple products (Gallego and Van Ryzin 1997), unknown market responses (Araman and Caldentey 2009, Farias and Van Roy 2010), multiple offered prices (Aydin and Ziya 2009), time-varying demand (Cao et al. 2012), discounts and rebates (Aydin and Ziya 2008, Hu et al. 2017), and sales target constraints (Du et al. 2022). In these works, demand for products is modeled as a controlled Poisson process. As such, they too model their problem as Poisson intensity control problems and subsequently study the corresponding HJB equations to establish structural properties of the optimal policy.

There are two fundamental differences between our formulation and the ones encountered in revenue management (see Gallego and Topaloglu (2019) for a complete overview of the field). First, a chosen price impacts both the revenue accumulation and demand process in a different way than a chosen production rate impacts the revenue accumulation and deterioration process in our

problem. In dynamic pricing, revenue only accumulates at discrete demand arrival epochs, while in our problem, revenue is accumulated continuously. In the former, choosing a higher price leads to a lower demand rate and potentially more revenue if demand occurs, while in our problem, choosing a higher production rate induces a higher deterioration rate and instantaneously a higher revenue rate. Second, we deal with a boundary condition that models the maintenance cost levied at the end of the planning horizon. This boundary condition needs to be taken into account when adopting a production policy throughout the planning horizon to prevent high maintenance costs at the end. In the revenue management literature, such boundary conditions typically do not exist.

3. Problem Formulation

We consider a single-unit production system whose deterioration can be represented by a single parameter. The deterioration process of the system is continuously monitored and described by a Poisson process $Y = (Y_t)_{t \geq 0}$, with $Y_0 = 0$ because at each planned maintenance moment a new unit is installed and hence has zero deterioration. Poisson deterioration is a common modeling choice in the OM literature, for recent examples, see, e.g., Tian et al. (2021) and Li and Tomlin (2022). The system fails when the deterioration level is equal to a fixed failure level $\xi \in \mathbb{N}$. This is in line with the prevailing maintenance literature and practice, see, e.g., Wang (2002), Gebraeel et al. (2005), Elwany et al. (2011), Uit het Broek et al. (2020), and Drent et al. (2023a), and the case studies therein. We denote the set of all deterioration levels with $\mathscr{X} \triangleq \{0, 1, \ldots, \xi\}$. (Discrete deterioration states are common, e.g. Derman (1963), Kim and Makis (2013), Tian et al. (2021), Li and Tomlin (2022), Drent et al. (2023a) and Zhang and Zhang (2023).)

The DM can adjust a univariate production control variable of the system that will both impact the deterioration process and the rate at which the production system generates revenue. We remark that in the remainder of the paper, we will refer to this variable as the production rate, though in practice, it could represent any adjustable setting or control variable of the system that has the said impact. Hence, if the DM selects production rate $s \in \mathscr{S} \triangleq [0, \bar{s}]$ (where $0 < \bar{s} < \infty$ is the maximum production rate), then the rate of the Poisson process is instantaneously governed by the base rate, denoted with λ, multiplied by the deterioration function $f : \mathscr{S} \to \mathbb{R}_+$, and the revenue rate is instantaneously governed by the revenue function $r : \mathscr{S} \to \mathbb{R}$. In practice, one can view the base rate as the base line for deterioration under normal operating conditions.

For now we shall assume that the base rate λ is fully known to the DM and that each system has the same fixed base rate upon replacement at the scheduled maintenance moment. The results for such systems with fixed base rates also help us to develop and benchmark a heuristic for systems for which the base rates are unknown a-priori to the decision-maker and differ from system to system. We discuss such populations of systems with unit-to-unit variability in detail in §6.

Throughout this paper, we use increasing and decreasing in the weak sense. We impose the following assumption on the deterioration and revenue function.

ASSUMPTION 1. *The deterioration function $f(\cdot)$ and revenue function $r(\cdot)$ are both continuous and increasing. Furthermore, $f(0) = 0$ and $r(0) \leq 0$.*

Assumption 1 together with the compactness of $\mathscr{S}$ also guarantees that $f(\cdot)$ and $r(\cdot)$ are bounded. The increasing assumption is reasonable in practice as higher production rates generally lead to higher revenue and deterioration rates. Two remarks regarding $r(\cdot)$ are in order here.

REMARK 1. Our analysis does not require $r(\cdot)$ to be increasing: All results also hold if $r(\cdot)$ is decreasing. However, if $r(\cdot)$ is decreasing then this will result in a practically meaningless policy in which the system is always turned off as turning it on leads to a decrease in revenues.

REMARK 2. In settings where the focus is to maximize production revenues, it is natural to have $r(0) = 0$ (no revenue is earned when the system is off), though there are situations where $r(0) < 0$. One practical setting is when a DM aims to satisfy a (continuous) demand rate, say $D > 0$, with the production rate s. This can be modeled by choosing $r(\cdot)$ such that the downfall $(D - s)^+$ is penalized and the surplus $(s - D)^+$ is rewarded, where $(x)^+ \triangleq \max(0, x)$, so that $r(0)$ is negative. One notable example is $r(s) \triangleq b(s - D)^+ - p(D - s)^+$, where b (p) is the bonus (penalty) rate. One extension to this case, which we discuss in Appendix B, is when a single demand rate D can be met by the cumulative production rate of multiple systems.

Maintenance is only performed at scheduled maintenance moments, and the next maintenance action is scheduled at time T (e.g., maintenance teams visit offshore wind farms semi-annually, or the system undergoes intentional shutdowns yearly). The maintenance cost at scheduled maintenance moments is governed by an increasing convex maintenance cost function $c_m : \mathscr{X} \to \mathbb{R}_+$, which reflects that higher deterioration levels lead to greater adverse effects, and these effects worsen as deterioration progresses. The canonical maintenance cost structure with a single corrective (c_u) and single preventive maintenance cost ($c_p < c_u$) belongs to this class of maintenance cost functions.

When maintenance moments are scheduled, the DM's operational problem is to dynamically adjust the production rate with the goal of maximizing the expected profit (i.e. revenue minus costs). We formulate this optimization problem with time indexed backwards. To this end, let $X_t = \{Y_{\tilde{t}} \mid 0 \leq \tilde{t} \leq T\}$ be the deterioration level at time $t = T - \tilde{t}$ before the next maintenance moment. Since $Y_{\tilde{t}}$ is the number of deterioration increments up to time $\tilde{t}$, $dY_{\tilde{t}} = 1$ if deterioration increases at time $\tilde{t}$. Let $\mathscr{U}$ be the set of all admissible non-anticipating production policies satisfying $\int_{k=0}^{t} dY_k \leq \xi - X_t$ (almost surely). This constraint ensures that whenever the system fails (i.e. when $Y_t = \xi$), the system is turned off. It can always be satisfied because $s = 0 \in \mathscr{S}$ with $f(0) = 0$ turns

the system off. For a given production policy $u \in \mathscr{U}$, let $s_\lambda^u(x, k)$ be the induced production rate if the deterioration level is x, the time to planned maintenance moment is $k \in [t, 0)$ and the base rate is λ. The expected profit, starting from state (x, t) under a policy $u \in \mathscr{U}$ is then denoted by

$$J_\lambda^u(x, t) \triangleq \mathbb{E}_u \left[\int_0^t r\big(s_\lambda^u(X_k, k)\big) dk - c_m\big(X_0\big) \;\Big|\; X_t = x \right], \tag{1}$$

where $\mathbb{E}_u$ denotes that the expectation is taken with respect to policy u. The integral part in (1) represents the expected revenue, while the second part represents the expected maintenance costs under policy u. We make the dependence of the total expected profit on the base rate λ explicit by using the subscript, thereby establishing the correspondence to a population of systems with the same fixed rate. The goal is to find the optimal condition-based production policy, denoted with u_λ^*, that maximizes the expected profit generated over $[t, 0]$, denoted with $J_\lambda^*(x, t) \triangleq \sup_{u \in \mathscr{U}} J_\lambda^u(x, t)$. For brevity, we introduce the shorthand notation $\Delta_\lambda(x, t) \triangleq J_\lambda^*(x, t) - J_\lambda^*(x + 1, t)$, which can be interpreted as the marginal value of less deterioration for fixed t and λ under the optimal policy. Using this additional notation, we are in the position to present our first result in Lemma 1 below: The HJB equation for the optimal expected profit $J_\lambda^*(x, t)$.

LEMMA 1. $J_\lambda^*(x, t)$ *is the solution of*

$$\frac{\partial J_\lambda^*(x, t)}{\partial t} = \begin{cases} \max_{s \in \mathscr{S}} \Big[r(s) - \lambda f(s) \Delta_\lambda(x, t) \Big], & \text{if } x \in \{0, 1, \ldots, \xi - 1\}, \\ 0, & \text{if } x = \xi, \end{cases} \tag{2}$$

with boundary conditions $J_\lambda^*(x, 0) \triangleq -c_m(x)$ *for all* $x \in \mathscr{X}$.

The formulation in Lemma 1 is a Markovian intensity control problem; we refer the reader to Bremaud (1980) for an excellent treatment of the general theory on this subject. Since $f(\cdot)$ and $r(\cdot)$ are continuous and bounded (by Assumption 1), and $\mathscr{S}$ is a compact subset of $\mathbb{R}$, we are guaranteed by Theorem III of Bremaud (1980) that there exists a solution to (2), which we denote by $s_\lambda^*(x, t)$; that is, $s_\lambda^*(x, t)$ is the optimal production rate for state (x, t) when the base rate is λ.

4. Fixed Base Rate: Analytical Results

This section contains our analytical results for systems with a known, fixed base rate. We first present structural properties of the optimal expected profit and optimal condition-based production policy for a given maintenance interval length T. We then study the tactical problem of scheduling maintenance moments, given that the optimal production policy is used in between such moments.

4.1. The Optimal Condition-Based Production Policy

We first show that the optimal expected profit function $J_\lambda^*(x,t)$ behaves monotonically in its state variables and the base rate, which we then use to prove that the optimal production rate also has certain monotonicity properties. We illustrate these theoretical results with numerical examples.

Using sample path arguments, one can deduce first-order monotonic properties of $J_\lambda^*(x,t)$ with respect to x, t, and λ that are intuitive and useful for subsequent analysis. These properties are summarized in the following lemma.

LEMMA 2. *The optimal expected profit $J_\lambda^*(x,t)$ is decreasing in the deterioration level x, increasing in the time to the next maintenance moment t, and decreasing in the base rate λ.*

Everything else fixed, Lemma 2 states that for a more deteriorated system, the expected profit will be lower than for a system that is less deteriorated. Similarly, everything else fixed, the expected profit is higher when the time to planned maintenance is higher, because there is more time to generate revenue. Finally, a system whose base rate is higher, deteriorates faster under any production rate, which in turn leads to a lower expected profit.

We now turn our attention to higher-order monotonic properties of the expected profit with respect to its parameters, which are needed to prove structural properties of the optimal policy.

PROPOSITION 1. *The optimal expected profit $J_\lambda^*(x,t)$ is concave in x for all $t \geq 0$ and $\lambda \geq 0$.*

The concavity of the optimal expected profit in the deterioration level implies that expected profit drops ever steeper as the system deterioration level is closer to the failure level for any fixed time. This is intuitive: A system that is approaching the failure threshold has less and less room (in terms of remaining useful condition) to control the deterioration, and hence for generating revenue.

The next result establishes that the marginal value of less deterioration, increases in the remaining time until the next planned maintenance.

PROPOSITION 2. $\Delta_\lambda(x,t)$ *is increasing in t for all $x \in \mathscr{X}$ and $\lambda \geq 0$.*

The intuition behind Proposition 2 is as follows. Consider a system with a better condition (i.e. with one deterioration level less) than another system. The former generates more revenue than the latter, see Lemma 2. Now, if the scheduled maintenance moment is further away, then for the former system, there is a longer planning horizon to exploit this better condition – for generating revenue – than when there is less time to do so. This leads to a higher additional revenue.

Using a similar intuition one can also reason that the optimal expected profit must be concave in the time to the next scheduled maintenance moment for all deterioration levels and each base rate. This is indeed true and we present this result in the next proposition.

PROPOSITION 3. *The optimal expected profit $J_\lambda^*(x,t)$ is concave in t for all $x \in \mathscr{X}$ and $\lambda \geq 0$.*

We now proceed with proving that the optimal expected profit is submodular in (x, λ), which subsequently enables us to establish the monotonic behavior of the optimal policy in the base rate λ in light of the well-known Topkis' Theorem (see, for instance, Topkis 1998, Theorem 2.8.1). We first state two technical results that relate to preservation of submodularity. These results are preceded by a formal definition of submodularity.

DEFINITION 1 (SUBMODULARITY). A bivariate function, $g(x,y)$, is submodular in (x,y) if for all $x^+ \geq x^-$ and $y^+ \geq y^-$, we have $\left(g(x^+, y^-) - g(x^-, y^-)\right) \geq \left(g(x^+, y^+) - g(x^-, y^+)\right)$.

LEMMA 3. *Suppose $g(y)$ is nonnegative, and increasing in y, and $f(x,y)$ is nonnegative, submodular in (x,y), and decreasing in x, then $g(y)f(x,y)$ is submodular in (x,y).*

LEMMA 4. *Suppose $f(x,\theta)$ is submodular in (x,θ) and concave in x and the nonnegative random variable $Z(\theta)$ is stochastically increasing in the usual stochastic order in θ, then $\mathbb{E}[f(x + Z(\theta),\theta)]$ is submodular in (x,θ) and concave in x.*

Lemma 3 and 4 complement other preservation results – usually needed for establishing structural properties in optimization problems – that can be found in the literature; for instance a closely related result by Chao et al. (2009) that focuses on $\mathbb{E}[f(x, Z(\theta))]$ (in our notation). Using these two lemma's, we are able to establish the submodularity of the optimal expected profit in (x, λ), which is presented in the result below.

PROPOSITION 4. *For all $t \geq 0$, the optimal expected profit $J_\lambda(x,t)$ is submodular in (x, λ).*

Definition 1 shows that submodularity of a bivariate function $g(x,y)$ in (x,y) implies that the difference $g(x^+, y) - g(x^-, y)$ is decreasing with respect to y. Using this definition, Proposition 4 thus implies that the decrease in the optimal expected profit due to more deterioration, i.e. $J_\lambda(x^+,t) - J_\lambda(x^-,t)$ with $x^+ \geq x^-$, is decreasing in the base rate. In other words, the decrease in the optimal expected profit due to deterioration accumulation is smaller for systems facing a lower base rate than for systems facing a higher base rate.

The next result is the main result of this section; it characterizes the monotonic behavior of the optimal production rate.

THEOREM 1. *The optimal production rate when the deterioration level is x, the time until planned maintenance is t and the base rate is λ, i.e. $s_\lambda^*(x,t)$, is decreasing in x, t, and λ.*

The intuition behind Theorem 1 is as follows.

1. **Monotonicity in x:** Given the same time until planned maintenance, production will be slowed down whenever failure is nearing to avoid high maintenance costs. Likewise, a DM will select a higher production rate for a system that is less deteriorated than for a system that is close to failure. In doing so, she increases revenues without running a high risk of failure.

2. **Monotonicity in t:** Given the same deterioration level, a DM is more conservative at the beginning of the planning horizon and chooses a lower production rate than when the scheduled maintenance moment is nearby. Similarly, when the planned maintenance moment is nearing and the deterioration level is kept constant, the DM will exploit the remaining useful condition for accumulating revenue by choosing an increasingly higher production rate.

3. **Monotonicity in λ:** If a system deteriorates faster under any production rate through a higher base rate, then a DM will choose lower production rates for all deterioration levels and remaining times than for a system that deteriorates more slowly. This ensures that in the former, high maintenance costs are prevented, while in the latter, more revenue is generated.

Notably, the structure of the optimal condition-based production policy is inherently different from the optimal policy under deterministic deterioration. For the latter, Uit het Broek et al. (2020) prove that it may be optimal to purposely break down the system when a failure is unavoidable, breaking down any monotonic properties of the optimal policy. This sharply contrasts with our result, where a DM would gradually decrease production when a potential failure is nearing. However, Theorem 1 does formally confirm some of their reported intuitions – based on numerical explorations – of the structure of the optimal policy for the stochastic setting (see Uit het Broek et al. 2020, §5.3).

The monotonic behavior of the optimal condition-based production policy is illustrated by Figures 1 and 2. In both figures, we use $\xi = 10$ (failure threshold), $c_m(\xi) = 5$ (corrective maintenance cost), and $c_m(\cdot) = 1$ for all non-failed states (preventive maintenance cost). The planning horizon T is equal to 15 and the production rate space is equal to $\mathscr{S} = [0,1]$. For the revenue and deterioration functions, we use $r(s) = s^{1/2}$ and $f(s) = s^2$. We obtained the optimal policy in these illustrations, and elsewhere in this paper, using the standard approach for solving continuous time dynamic programs, based on the finite difference formulation (see Appendix C for more details).

In the left (right) subfigure of Figure 1 the base rate is 1 (4). In each subfigure, we clearly observe that the optimal production rate increases as the time to scheduled maintenance decreases for a fixed deterioration level (moving right horizontally), and that the optimal production rate decreases as the condition deteriorates (moving up vertically). If we compare the left with the right subfigure, we clearly see that lower rates are prescribed in all states (x, t) in the right subfigure.

Figure 2 presents sample paths of the optimal condition-based production policy (bottom) and the corresponding controlled deterioration (top). We see that when the system's condition deteriorates, production is immediately lowered, which highlights the monotonicity of the optimal
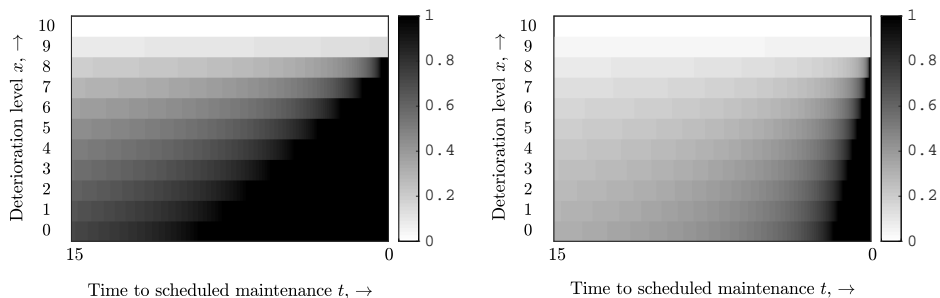
Figure 1   Illustration of monotonic behavior of the optimal condition-based production policy with base rate $\lambda = 1$ (left) and base rate $\lambda = 4$ (right). The color-bar indicates the optimal production rate.
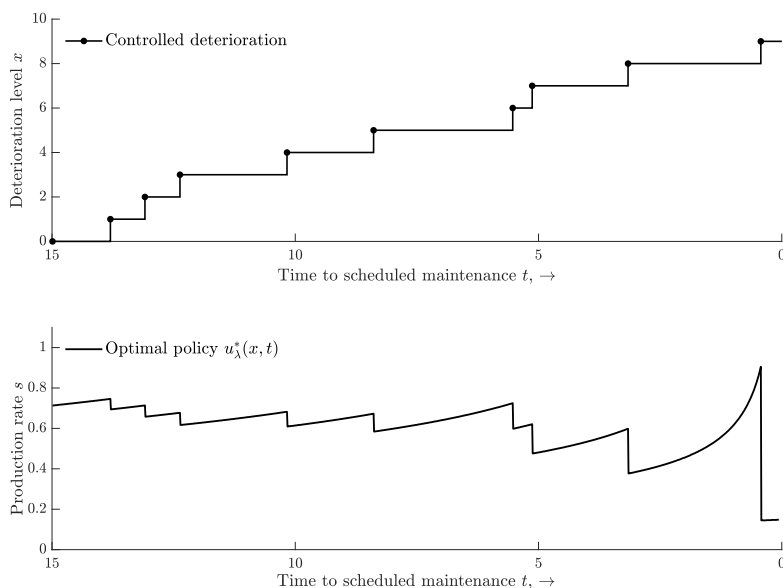

Figure 2   Sample paths of the optimal condition-based production policy (bottom) and controlled deterioration process (top), when the base rate λ is 1.

production rate in the deterioration level for fixed time to scheduled maintenance. Also, the gradual rise of the production rate between deterioration increments highlights the monotonicity of the optimal production rate in the time to scheduled maintenance for a fixed deterioration level.

4.1.1. Optimality of Bang-Bang Production Policies   In this subsection, we present sufficient conditions for the optimal policy to be a so-called bang-bang policy. See Definition 2 below for a formal description of a bang-bang policy.

DEFINITION 2 (BANG-BANG POLICY). A bang-bang production policy is a policy in which production is either set to the maximum rate $s = \bar{s}$ (full production) or turned off $s = 0$ (no production).

The next result provides a sufficient condition in terms of $r(\cdot)$, $f(\cdot)$, and their maxima, for a bang-bang policy to be optimal. Note that this condition does not depend on the base rate λ.

PROPOSITION 5. *The optimal policy has a bang-bang structure if*

$$\frac{r(s)}{f(s)} \leq \frac{r(\bar{s})}{f(\bar{s})} \ for \ all \ s \in (0, \bar{s}). \tag{3}$$

The next result, which follows from Proposition 5, provides sufficient conditions only on the functions $r(\cdot)$ and $f(\cdot)$ for a bang-bang policy to be optimal that can be verified without verifying whether (3) holds for all $s \in (0, \bar{s})$.

COROLLARY 1. *The optimal policy has a bang-bang structure if any of the following conditions is satisfied:*

  (i) $r(s) = f(s)$ *for all* $s \in \mathscr{S}$.
 (ii) $r(s) = a \cdot s^{\alpha}$ *and* $f(s) = b \cdot s^{\beta}$ *with constants* $\alpha \geq \beta$, a, *and* b.
(iii) $r(\cdot)$ *is convex and* $f(\cdot)$ *is concave.*

We remark that Corollary 1 (*iii*) provides theoretical foundation for numerical observations of Uit het Broek et al. (2020, p. 804). Their experiments indicate that bang-bang policies – or, on-off policies in their nomenclature – have the same expected profit as the optimal policy under stochastic deterioration when the deterioration function is concave and the revenue function linear.

Under a bang-bang policy, the policy itself can be described by a switching curve, denoted with $\delta^{(\lambda, t)}$, such that the system, and hence production, is switched off if the deterioration level is greater than or equal to this curve. The following result establishes the monotonic behavior of this switching curve. This result is an immediate consequence of Theorem 1; hence we omit the proof.

COROLLARY 2. *For a bang-bang policy, the switching curve* $\delta^{(\lambda, t)}$ *is decreasing in both the time to scheduled maintenance t and the base rate* λ.

The optimality of a monotonic bang-bang policy is not only convenient because of its practical appeal but it also significantly decreases the computational burden of computing the optimal policy. Indeed, it effectively reduces the action space of our decision problem from a compact set with infinitely many actions to a discrete set with only two actions.

Figure 3 illustrates the bang-bang policy. In this example we use the same values as in Figure 1, except that we swap the revenue and deterioration function to comply with Corollary 1, i.e. $r(s) = s^2$ and $f(s) = s^{1/2}$. In line with Corollary 2, we see that the switching curve is decreasing in the time to scheduled maintenance, and decreasing in the base rate.

REMARK 3. The optimal policy when multiple production systems seek to meet a joint demand rate is explored numerically in Appendix B. Here we observe that different production systems synchronize their production rates to meet the demand rate with machines in better health carrying a higher proportion of the total production rate.
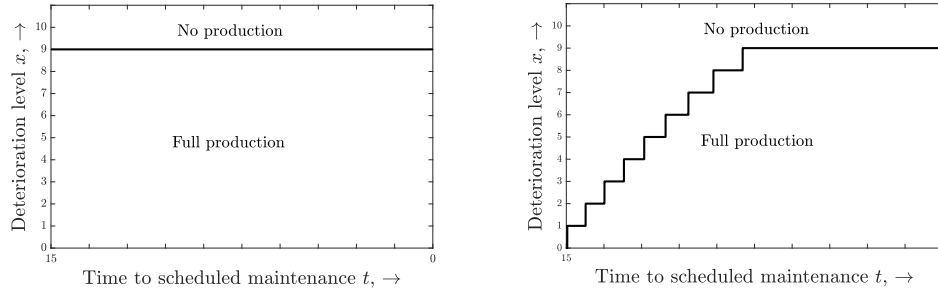
Figure 3   Illustration of a bang-bang policy for base rate $\lambda = 1$ (left) and base rate $\lambda = 4$ (right). The system is
switched off if the deterioration level is greater than or equal to the switching curve (i.e. black line).

4.2. Optimal Integration of Maintenance and Condition-Based Production

Theorem 1 characterizes the structure of the optimal production policy for the operational decision
problem given that maintenance is performed every T time units. In practice, a DM also needs
to decide upon the value of T in which she deals with a tactical trade-off. Specifically, frequent
planned maintenance (small T) leads to unnecessary maintenance activities and associated costs,
while infrequent maintenance (large T) both increases system failure risk and reduces revenue due
to necessitated lower production rates. In deciding upon T, a DM seeks to maximize the average
profit, by incorporating both the operational and tactical trade-off. More formally, let $g(T)$ denote
the average profit per time unit when an optimal condition-based production policy u_λ^* is followed
on the operational level and the planned maintenance moment is scheduled every T time units,
that is, $g(T) \triangleq \frac{1}{T} J_\lambda^*(0, T)$. The DM seeks to find $T^* \triangleq \arg\sup_{T \in \mathbb{R}_{>0}} g(T)$. This decision problem is
complex because of its nested structure. The next result simplifies its optimization considerably.

THEOREM 2. *Let T' be a strict local maximizer of $g(T)$, then T' is the strict global maximizer.*

Condition-based production on an operational level and maintenance decisions on a tactical level
should be made in conjunction rather than in isolation, and Theorem 2 provides us with a way
to do so. Specifically, readily available uni-modal optimization methods can be used to efficiently
find the optimal length of the maintenance interval. Figure 4a illustrates such a uni-modal average
profit function $g(T)$. In this example, where we use the same settings as in Figures 1 and 2, the
optimal maintenance interval length is equal to 8.6 with an average profit of 0.84 per time unit.

In the next section, we present numerical studies for instances as they are typically encountered
in practice. For all these instances, $g(T)$ has a strict local maximizer and hence, by virtue of
Theorem 2, a unique maximizer. Unfortunately, this does not hold in general; we found that $g(T)$
may also be non-decreasing and negative for all T. Figure 4b illustrates this case, where $g(T)$ is
negative for all T and with $g(T)$ approaching zero from below as T grows large. In this example,
$\lambda = 1$ (base rate), $\xi = 10$ (failure level), $c_m(\xi) = 50$ (corrective maintenance cost), and $c_m(\cdot) = 40$
for all non-failed states (preventive maintenance cost). The production rate space is $\mathscr{S} = [0, 1]$ and
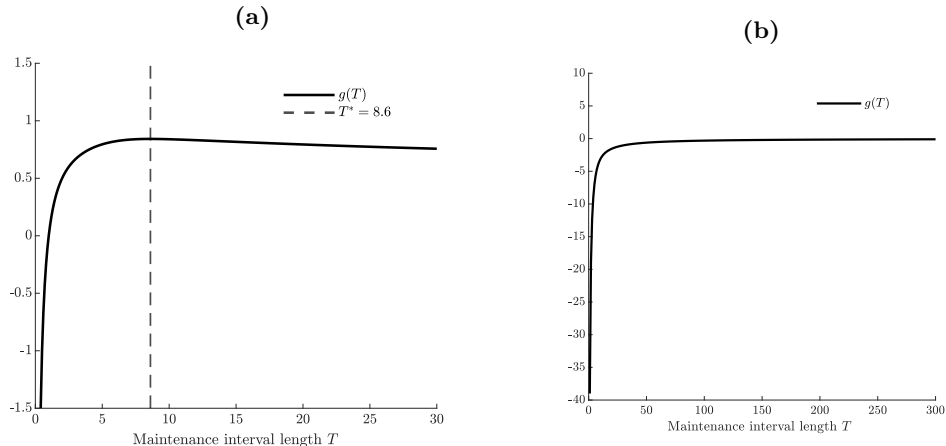
Figure 4   Two functions $g(T)$ when the base rate is 1. In Panel (a), the optimal maintenance interval is depicted with the dashed line. Panel (b) illustrates an example with no global maximizer.

we use $r(s) = s^2$ and $f(s) = s^{1/2}$ for the revenue and deterioration functions, respectively. Under this set of choices, the system deteriorates relatively faster than it generates revenue for all $s \in \mathscr{S}$, which, in combination with the high maintenance costs at the end of the planning horizon, results in a non-profitable system. In such a case, the potential revenue under any production policy is not sufficient to compensate the high maintenance costs, and choosing a large value of T ensures that at least the maintenance costs are on average very low per time unit. Such non-profitable systems are arguably not the ones we operate in practice.

5. Value of Condition-Based Production and Maintenance Management

This section contains two comprehensive numerical studies that highlight the practical value of our theoretical results on populations of systems with the same fixed base rate. The goal of this section is twofold. Firstly, we quantify the value of dynamically adjusting the production based on the system's deterioration and time to scheduled maintenance. Secondly, we investigate the value of integrating maintenance and condition-based production decisions. In both numerical studies, we use the canonical maintenance cost structure with a fixed preventive maintenance cost $c_m(\cdot) \triangleq c_p$ for all non-failed states, and a corrective maintenance cost $c_m(\xi) \triangleq c_u$ for the failed state.

5.1. Static versus Condition-Based Production

We compare the optimal condition-based production policy with the optimal static policy that cannot control production in real time. Under this static policy, the DM chooses a fixed production rate, which is operated continuously in between consecutive maintenance moments until either failure or until the maintenance moment, whichever comes first. We call this static policy the fixed-rate policy, denoted by π_{FS}, and obtain it by solving the following maximization problem:

$$\pi_{FS} \in \underset{s \in \mathscr{S}}{\arg\max} \left[\mathbb{E}\big[r(s) \cdot \min\{T_\xi, T\} \big] - c_p - (c_u - c_p) \cdot \mathbb{P}\big[T_\xi \leq T \big] \right], \qquad (4)$$

where T_ξ is an Erlang random variable with shape ξ and rate $\lambda f(s)$, and T is the length of the maintenance interval. The first part of the objective function is the expected accumulated revenue, while the second part is the expected maintenance cost at time T. The maximization problem can be easily solved numerically to obtain π_{FS}.

In order to assess the value of using the optimal condition-based production policy, π_{CS}, instead of the static fixed-rate policy, π_{FS}, we use the relative profit increase, denoted by $\mathscr{R} = 100 \cdot (P_{\pi_{CS}} - P_{\pi_{FS}})/P_{\pi_{FS}}$, where P_π is the expected profit (over the planning horizon) of policy $\pi \in \{\pi_{CS}, \pi_{FS}\}$.

5.1.1. Full factorial experiment In this section and Appendix D.1, we report on a full factorial experiment to quantify the value of condition-based production averaged across a wide range of parameter settings. Our test bed consists of 3750 instances, obtained by permuting various values for all input parameters. See Table 4 in Appendix D.1 for all input parameter values of the test bed; here we describe only a subset. We use the functions $f = s^\gamma$ and $r = s^\nu$ to represent the deterioration and revenue functions, respectively. We use the interval $\mathscr{S} = [0, 2]$ for the production rates to choose from. By using different values for γ and ν, our test bed consists of concave, linear, and convex deterioration and revenue functions. Furthermore, we use different base rates and failure levels so that we vary the expected time to failure under normal deterioration.

The main results of the full factorial experiment are summarized in Table 1, which presents the average and standard deviation of $\mathscr{R}$, as well as the average value $P_{\pi_{CS}}$ over all instances. Detailed results are provided in Appendix D.1. Table 1 indicates that adjusting the production rate based on the system's deterioration and time to maintenance, instead of using the best fixed-rate policy, leads on average to sizeable profit increases across a wide range of parameter settings. These profit increases are mainly driven by two mechanisms. First, the condition-based production policy is able to prevent failures by slowing down deterioration when the failure threshold is nearing, thereby saving on high corrective maintenance costs. This is different from the fixed-rate policy, which just continues to operate with the fixed production rate until failure. Second, when the scheduled maintenance moment is nearing and the system is still in good condition, the condition-based production policy will exploit the remaining useful condition to accumulate revenue by setting a higher production rate. Such opportunities are not seized by fixed-rate policies.

Table 1 Main results static versus condition-based production.

Mean $\mathscr{R}$	SD $\mathscr{R}$	$P_{\pi_{CS}}$
50.42	202.46	21.38

While on average condition-based production leads to significant profit increases, Table 5 in Appendix D.1 also shows that the exact value depends on the specific input parameters. The value

seems to depend strongly on the expected time to failure under normal deterioration (through λ and ξ) and on how production affects both deterioration and revenue accumulation (through ν and γ). In the next section, we investigate the impact of these input parameters on the value of condition-based production through comparative static analyses.

5.1.2. Comparative statics   We define a base case of parameter settings, and then study how parametric changes with respect to this base case affect the relative profit increase of condition-based production $\mathscr{R}$. For the base case, we set the base rate $\lambda = 1$, the failure level $\xi = 14$, the preventive maintenance cost $c_p = 2$, the corrective maintenance cost $c_u = 10$, $\gamma = \nu = 1$, and the planning horizon $T = 10$. Note that these parameters are set such that they coincide with their respective median values of the full factorial experiment above, except for the planning horizon T.

We first investigate how the relative value of condition based production is affected by the expected time to failure under normal deterioration. Figure 5a illustrates that as the expected time to failure increases (i.e. starting from $(\lambda, \xi) = (1.4, 10)$ moving along either axis), the value of optimal condition based production polices over static fixed-rate policies decreases substantially from more than 180% to less than 30%. Comparing Figures 5b and 5c, we see that the optimal condition-based production policy consistently accumulates more revenue than the optimal static fixed-rate production policy. Observe further that this difference diminishes as the expected time to failure increases, while the absolute revenues of both policies increase owing to the increased time available for production. Figure 5e indicates that the optimal rate of the fixed-rate production policy increases in the expected time to failure. This enables more revenue accumulation, but comes at the expense of a higher probability of corrective maintenance at the end of the planning horizon, as indicated in Figure 5e, and thus a higher total maintenance cost, as indicated in Figure 5f. By contrast, under the optimal condition-based production policy, the deterioration level never reaches the failure threshold for any of the instances considered in this analysis, resulting in only incurring the preventive maintenance cost at the end of the planning horizon.

Figure 5 suggest – and additional numerical experiments confirm this – that as the expected time to failure increases, the optimal rate of the fixed-rate policy increases towards the maximum production rate. For sufficiently large expected times to failure, the probability of corrective maintenance goes to zero under the maximum production rate, thus implying that the revenue and cost of both the fixed-rate and the condition-based production policy converge to the same values. That is to say, the relative profit increase of condition-based production over fixed-rate production goes to zero as the expected time to failure grows large. This argument can be made more rigorous by noting that the probability of corrective maintenance $\mathbb{P}[T_\xi \leq T]$ in the expected profit function (4) under the fixed rate $\bar{s}$ ($< \infty$) goes to zero as the expected time to failure $\mathbb{E}[T_\xi]$ grows large, through increasing the failure level ξ and/or decreasing the base rate λ.
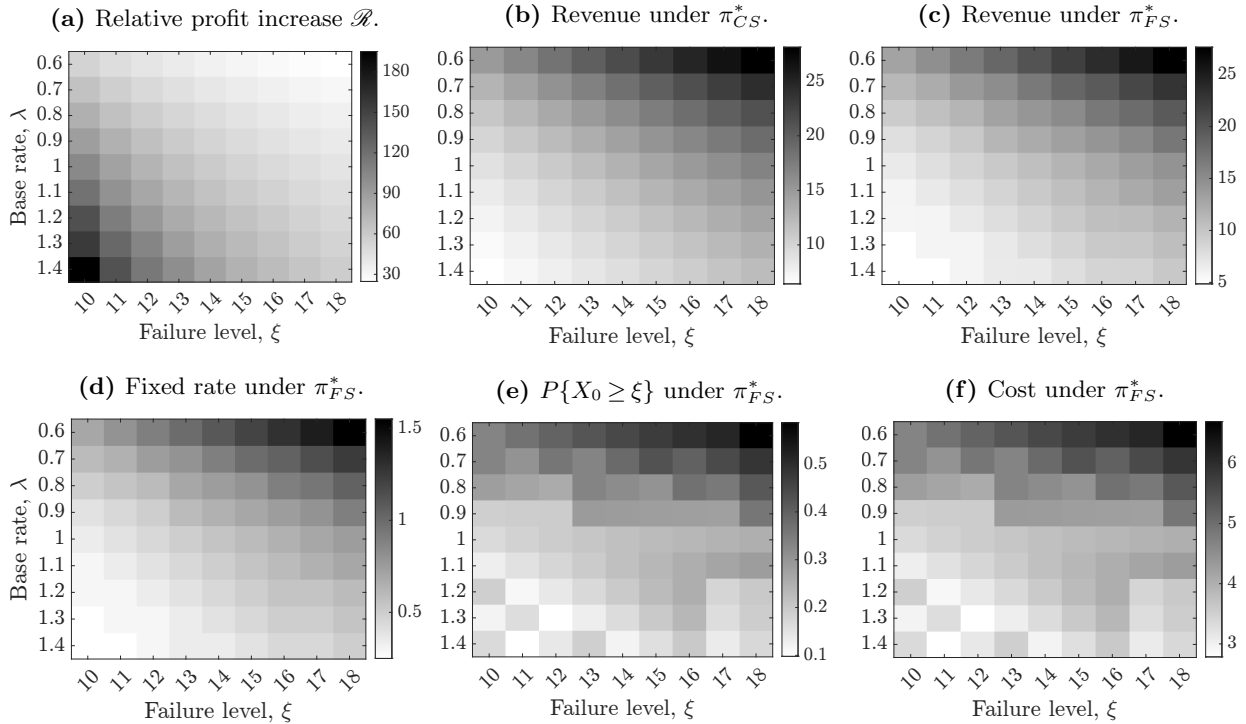
(a) Relative profit increase $\mathscr{R}$.   **(b)** Revenue under π^*_{CS}.   **(c)** Revenue under π^*_{FS}.

(d) Fixed rate under π^*_{FS}.   **(e)** $P\{X_0 \geq \xi\}$ under π^*_{FS}.   **(f)** Cost under π^*_{FS}.

Figure 5   Various output measures for different values of the base rate λ and the failure level ξ when the other parameters are held fixed at the base case values. Note that the expected time to failure under normal deterioration increases as the base rate (failure level) decreases (increases).

Next we investigate the effect of the deterioration and revenue functions on the value of condition-based production. Figure 6a indicates that the relative profit increase of optimal condition-based production over static fixed-rate production increases substantially from less than 20% to more than 120% as we move closer to the anti-diagonal consisting of instances for which the revenue function and deterioration function parameter are identical. This suggests that condition-based production is most beneficial in settings where production impact on revenue and deterioration are similar. Figure 6d indicates that the optimal fixed rate equals $\bar{s}$ for most of the instances left of the anti-diagonal where the optimal condition-based production policy is of the bang-bang type (i.e. $\nu \geq \gamma$). This leads to slightly higher revenue accumulation than what the optimal condition-based production policy accumulates. However, different from the condition-based production policy, the fixed-rate policy cannot be turned off when failure is nearing. Consequently, the fixed-rate policy almost always leads to failure in these instances, as indicated by the costs of these instances in Figure 6f being roughly equal to the corrective maintenance cost. For all instances on the right of the anti-diagonal (i.e. $\gamma > \nu$) we observed numerically that the optimal condition-based production policy is not of the bang-bang type. For these instances, we deduce that condition-based production consistently generates higher revenues than fixed-rate production while simultaneously incurring lower maintenance costs, and that absolute revenues and profits of these instances under both production policies are typically lower than the instances for which $\nu \geq \gamma$.
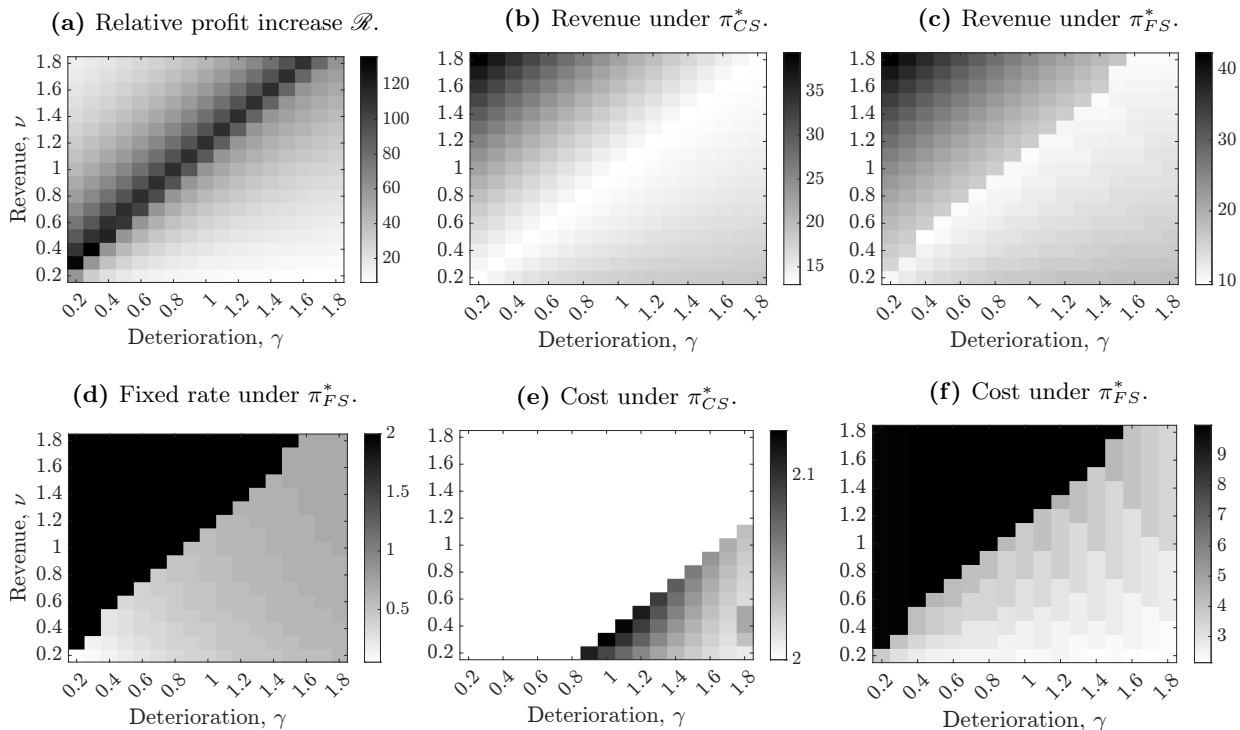
(a) Relative profit increase $\mathscr{R}$.   **(b)** Revenue under π_{CS}^*.   **(c)** Revenue under π_{FS}^*.

(d) Fixed rate under π_{FS}^*.   **(e)** Cost under π_{CS}^*.   **(f)** Cost under π_{FS}^*.

Figure 6   Various output measures for different values of the revenue function parameter ν and the deterioration function parameter γ when the other parameters are held fixed at the base case values. Note that the optimal condition-based production policy for the instances on the anti-diagonal and above (i.e. $\nu \geq \gamma$) is of the bang-bang type (cf. Corollary 1).

Finally, Table 5 in Appendix D.1 indicates that the relative profit increase of condition-based production averaged across a wide range of settings as well as the variability in this profit increase across those settings increase (decrease) in ν (γ). Our discussions above provides explanation for this behavior: As ν (γ) increases (decreases), the number of instances for which the optimal condition-based production policy has the bang-bang structure increases (decreases) as well (i.e. $\nu \geq \gamma$), and our observations above become more persistent.

5.2. Sequential versus Integrated Maintenance Scheduling

Theorem 2 enables efficient optimization of the maintenance interval length when the optimal condition-based production policy is followed in between those maintenance moments. This allows DMs to jointly decide upon the two policies (i.e. maintenance-interval and production policies). In practice, these two decisions are often made sequentially: First, the tactical decision on the maintenance interval length is made, followed by the operational decisions on production control based on that interval length. This is typically approached as follows. The maintenance interval length problem – without taking into account the production policy and using only the knowledge of the base rate – is a classical age-based maintenance problem (Barlow and Hunter 1960). The

lifetime under the base rate is an Erlang random variable with shape ξ and rate λ. Hence, the average maintenance cost rate for a given maintenance interval length t, denoted $g_{mc}(t)$, equals

$$g_{mc}(t) = \frac{c_p + (c_u - c_p) \cdot \mathbb{P}\big[T_\xi \leq t\big]}{\mathbb{E}\big[\min\{T_\xi, t\}\big]}, \tag{5}$$

where the numerator is the expected maintenance cost, and the denominator is the expected time to maintenance (either preventive or corrective). Since the hazard rate of an Erlang random variable is increasing, we know that the function $g_{mc}(t)$ has a unique minimizer that is easily computed. Given this optimized maintenance interval length, denoted t_S^*, we can then use the optimal condition-based production policy to control the system's deterioration in between those maintenance moments, which yields an expected profit rate per time unit. We call this approach the sequential approach, denoted with π_S. The resulting average profit rate is defined as $\hat{P}_{\pi_S} \triangleq J_\lambda^*(0, t_S^*)/t_S^*$.

Alternatively we can use Theorem 2 and determine the optimal length of the maintenance interval when taking into account that the optimal condition-based condition-based production policy is followed in between those maintenance moments. We call this approach the integrated approach, denoted with π_I. The resulting average profit rate is given by $\hat{P}_{\pi_I} \triangleq J_\lambda^*(0, t_I^*)/t_I^*$, where t_I^* is the optimal length of the maintenance interval, readily obtained by virtue of Theorem 2. The relative value of integrating maintenance and production policies is expressed by $\hat{\mathscr{R}} = 100 \cdot (\hat{P}_{\pi_I} - \hat{P}_{\pi_S})/\hat{P}_{\pi_S}$.

5.2.1. Full factorial experiment As before, we first report on a full factorial experiment to quantify the value of integrating maintenance and condition-based production averaged across a wide range of settings. We use the same parameters as the previous full factorial experiment except the parameter T, resulting in 1875 instances. The main results of this full factorial experiment are summarized in Table 2, which shows the average and standard deviation of $\hat{\mathscr{R}}$, as well as the average value $\hat{P}_{\pi_I}$ over all instances. Detailed results can be found in Appendix D.2.

Table 2 Main results sequential versus integrated maintenance scheduling.

Mean $\hat{\mathscr{R}}$	SD $\hat{\mathscr{R}}$	$\hat{P}_{\pi_I}$
21.39	28.35	1.83

Table 2 shows that integrating the length of maintenance intervals with the optimal condition-based production policy for revenue accumulation in between those maintenance moments leads to significant profit rate increases as opposed to treating them sequentially, with an average increase of 21.39%. Given that both approaches apply the optimal condition-based production policy on the operational level, we emphasize that this average profit increase can be fully attributed to the integration of maintenance and condition-based production. As with the value of condition-based production explored in the previous section, Table 6 in Appendix D.2 indicates that the value of the integrated approach is significantly influenced by the expected time to failure and how production affects both deterioration and revenue accumulation. We will now investigate this further.

5.2.2. Comparative statics We follow the same approach as in the comparative static analyses of the previous section. The base case remains the same, with the input parameters set such that they coincide with their respective median values used in the full factorial experiment (except T for obvious reasons). We note that for all instances in this section, we observed numerically that the optimal condition-based production policy on the operational level ensures that the system never fails under both the sequential and the integrated approach. Consequently, only the preventive maintenance cost is incurred upon the scheduled maintenance moment for all instances.
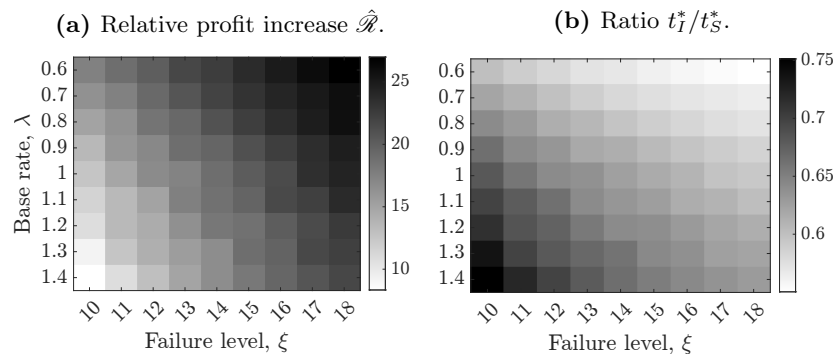

(a) Relative profit increase $\hat{\mathscr{R}}$. **(b)** Ratio t_I^*/t_S^*.

Figure 7 Output measures for different values of the base rate λ and the failure level ξ when the other parameters are held fixed at the base case values.

Figure 7a indicates that the value of integrating maintenance and condition-based production increases from less then 10% to more than 25% as the expected time to failure increases. While both approaches accumulate more profit per time unit as the expected time to failure increases, Figure 7a indicates that the integrated approach does so at an increasingly higher rate. This latter observation can be explained by means of Figure 7b. As the expected time to failure increases, both approaches increase the maintenance interval length, enabling more revenue accumulation on the operational level. However, as depicted in the figure, the maintenance interval length under the integrated approach grows slower than under the sequential approach. This highlights the advantages of integrating maintenance and condition-based production decisions as opposed to treating them sequentially. The integrated approach sets the maintenance interval length such that revenue accumulation in between those maintenance moments can be maximized while taking into account the maintenance cost at the end of the maintenance interval. By contrast, the sequential approach neglects this revenue accumulation when determining the maintenance interval, leading to longer intervals in which less revenue is accumulated to avoid the cost of failure replacement. The integrated approach's shorter maintenance intervals lead to more frequent maintenance moments and associated preventive maintenance costs. However, these higher costs are significantly offset by the increased revenue accumulation, especially for longer expected times to failure, all while

keeping the risk of corrective maintenance small. Thus the integrated approach strikes an optimal balance between increased maintenance costs and greater revenue generation.
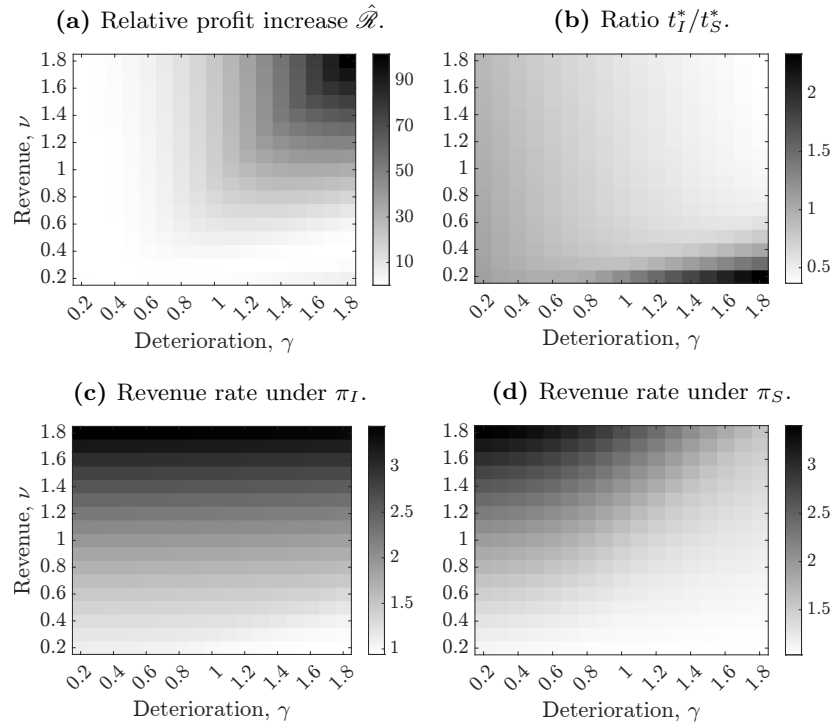
(a) Relative profit increase $\hat{\mathscr{R}}$.   **(b)** Ratio t_I^*/t_S^*.

(c) Revenue rate under π_I.   **(d)** Revenue rate under π_S.

Figure 8   Various output measures for different values of the revenue function parameter ν and the deterioration function parameter γ when the other parameters are held fixed at the base case values.

Next we investigate the impact of the deterioration and revenue functions on the value of integrating maintenance and condition-based production decisions. Figure 8a indicates that this value increases as the impact of production on revenue accumulation and/or deterioration becomes increasingly higher through increases in ν and/or γ. This is expected: The sequential approach neglects operational production decisions in between maintenance moments, and in particular their effects on revenue accumulation and deterioration. As these effects increase, the value of setting the maintenance interval length in conjunction with the operational production decisions increases. Figure 8b depicts the ratio between the optimal maintenance interval length under the integrated approach and the maintenance interval length under the sequential approach, the latter being constant for all instances as it assumes deterioration accumulation under normal operation (cf. Equation (5)). For the instances left on the anti-diagonal, where the optimal condition-based production policy is of the bang-bang structure, we see that the optimal maintenance interval length decreases as the impact of production deterioration increases. By contrast, the optimal maintenance interval length is robust against changes in the impact of production on revenue accumulation

since this impact does not affect the time to failure and only leads to more revenue accumulation (see Figure 8c). Producing at the maximum rate leads to faster deterioration under higher values of the deterioration function parameter γ. Hence if the optimal policy has a bang-bang structure, shortening the maintenance interval length allows for more revenue accumulation per time unit, outweighing the more frequently incurred preventive maintenance costs. We observe a different structure for the instances on the right of the anti-diagonal. If all production rates lead to more deterioration, lengthening the maintenance interval length allows for more conservative condition-based production rates such that revenues can be maximized without the risk of failure.

Figures 8c and 8d show that revenue rates increase as the revenue function parameter increases. By integrating maintenance length intervals and production decisions, Figure 8c clearly illustrates that this revenue rate is unaffected by increases in the deterioration function parameter. By contrast, Figure 8d shows that the revenue rate decreases in the deterioration function parameter under the sequential approach. This is due to the fact that the maintenance interval length is kept constant under the sequential approach, necessitating a condition-based production policy that aims to avoid failure, rather than to accumulate revenue.

6. Unit-To-Unit Variability: Learning the Base Rate

The previous sections focused on the situation in which the base rate is known and fixed for each system. In this section, we consider the situation in which the base rate is not known and where each newly installed system has a different base rate. In dealing with such unit-to-unit variability in base rates, a DM needs to learn the individual base rate based on the system's deterioration path and the used production policy. Such populations of systems subject to shock deterioration with variability in the base rates are quite prevalent in practice (see, e.g., Drent et al. 2023a).

We first propose a Bayesian procedure to learn the unknown base rate and use it to develop a heuristic policy that is inspired by the case where the base rate is known. We then discuss the results of a simulation study in which we compare the performance of the heuristic with an Oracle policy that knows the true base rate a-priori.

6.1. Bayesian Learning under Production Policies

In line with previous research on unit-to-unit variability in shock deterioration processes (e.g., Drent et al. 2023a,b), we model unit-to-unit variability in the systems' base rates with a Gamma distributed random variable, denoted by Λ, with known shape α_0 and rate β_0; that is, $\Lambda \sim \Gamma(\alpha_0, \beta_0)$. Each newly installed system has an a-priori unknown realization of Λ as base rate. Hence at $t = 0$ (T time units before planned maintenance), upon installation of a new system, the DM's knowledge regarding the unknown base rate is modeled by the prior density function $\varphi_0(\lambda | \alpha_0, \beta_0) =$

$\frac{\beta_0^{\alpha_0} \lambda^{\alpha_0-1} e^{-\beta_0 \lambda}}{\Gamma(\alpha_0)}$, where the subscript 0 indicates that this represents $t = 0$. Suppose that the DM adopts a production policy $\boldsymbol{s}_t \triangleq \{s_u\}_{u \in [0,t)}$ with production rate $s_u \in \mathscr{S}$ at time u over the period $[0, t)$ and the deterioration level at time t is $Y_t = y$, the proposition below shows that the posterior distribution at time t is then again a Gamma distribution, but with updated parameters.

PROPOSITION 6. *The updated posterior distribution of Λ at time t given realized deterioration level $Y_t = y$ and an employed production policy $\boldsymbol{s}_t$ up to time t, is a Gamma distribution with parameters $\alpha_0 + y$ and $\beta_0 + \int_{u=0}^{t} f(s_u) du$; that is, $\Lambda \sim \Gamma \left(\alpha_0 + y, \beta_0 + \int_{u=0}^{t} f(s_u) du \right)$.*

Proposition 6 implies that inference of the unknown base rate at each time t only requires the employed production policy $\boldsymbol{s}_t$ and the deterioration level $Y_t = y$. The mean of this updated Gamma distribution is equal to $\frac{\alpha_0 + y}{\beta_0 + \int_{u=0}^{t} f(s_u) du}$, which has intuitive appeal. If a system has been controlled by a policy with a higher aggregate amount of production rates captured by $\int_{u=0}^{t} f(s_u) du$ (recall that $f(\cdot)$ is increasing) than another, but both have the same deterioration, then we expect the base rate of the former to be smaller. Similarly, if a policy has been operated with the same aggregate amount of production rates, but one system has a higher deterioration level than another, then we expect the system with more deterioration to have a larger base rate.

6.2. A Certainty-Equivalent Policy

We now propose an intuitive heuristic policy called the certainty-equivalent (CE) policy to deal with unit-to-unit variability in the base rate that makes use of the Bayesian framework described in the previous section. Under this policy, the DM estimates the base rate based on her current knowledge and assumes that the true base rate is equivalent to her estimation with certainty. The CE policy is parameterized by a positive integer $N^{\mathrm{opt}} \in \mathbb{N}$, which denotes how often the DM re-optimizes the policy based on the accrued knowledge. In the CE policy, we divide the planning horizon in $N^{\mathrm{opt}} + 1$ phases, each of equal length.

At the start of the planning horizon, upon installation of the system, the DM's estimate for the base rate, denoted with $\hat{\lambda}_0$, is equal to the mean of the initial prior distribution: $\hat{\lambda}_0 = \frac{\alpha_0}{\beta_0}$. Using this estimate, we obtain the optimal policy for the known case with $\hat{\lambda}_0$ as base rate, and implement this policy until the first re-optimization epoch at $t_1 = \frac{T}{N^{\mathrm{opt}} + 1}$. At this point, the DM updates the estimate of the base rate based on the current deterioration level, denoted with x_1, and the policy used $\boldsymbol{s}_{t_1}$, to $\hat{\lambda}_1 = \frac{\alpha_0 + x_1}{\beta_0 + \boldsymbol{s}_{t_1}}$ and computes the optimal policy as if $\hat{\lambda}_1$ is the true base rate. This policy is then implemented until the next re-optimization epoch, after which the procedure is repeated until the scheduled maintenance moment. Figure 9 provides a schematic illustration of the CE policy with $N^{\mathrm{opt}} = 1$. Recall that u_λ^* denotes the optimal policy for a system with known base rate λ.

Compute $\hat{\lambda}_1 = \frac{\alpha_0 + x_1}{\beta_0 + s_{t_1}}$, and obtain policy $u^*_{\hat{\lambda}_1}$

Use policy $u^*_{\lambda_0}$ Use policy $u^*_{\hat{\lambda}_1}$

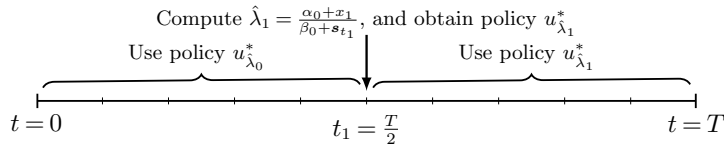$t = 0$ $t_1 = \frac{T}{2}$ $t = T$

Figure 9   Schematic illustration of CE policy with $N^{\mathrm{opt}} = 1$.

The CE policy is easy to implement and its behavior is well-understood from the structural properties established in Theorem 1. Specifically, if the DM learns that a system has a high (low) base rate, she will decrease (increase) the production in all states. The policy thus learns when it is possible to increase production to accumulate more revenue, but also when lower production is needed to prevent failures and high associated maintenance costs.

6.3. Performance of the Certainty-Equivalent Policy

In the numerical studies in §5 we were able to numerically compute the performance measures needed for the comparisons in those studies. The performance of the CE policy, however, can only be evaluated through simulation. In this section and Appendix D.3 we report on the results of a comprehensive simulation study in which we compare the performance of the CE policy with an Oracle policy that knows the true base rate upon installation.

We conduct a full factorial experiment in which we vary all input parameters across various values. The test bed of this experiment closely resembles the ones used in §5, with the addition of two input parameters that are specific to the setting studied here. First, in a simulation instance (we shortly describe the simulation procedure), the base rates are sampled from a Gamma distribution in which we fix the mean to 1, and vary its coefficient of variation, denoted with cv_Λ. This coefficient of variation is a measure for the unit-to-unit variability in the base rates of the systems. Second, the CE policy requires as input the parameter N^{opt}, which denotes how often the DM re-optimizes the policy based on the accrued knowledge. Permuting all input parameter values lead to a total of 4050 instances, see Table 7 in Appendix D.3 for a complete overview.

We benchmark the performance of the certainty equivalent policy, denoted with π_{CEP}, by the notion of relative regret, i.e. the percentage of expected profit loss (over a planning horizon) relative to the Oracle policy π_O. We compute this relative regret as $\bar{\mathcal{R}} = 100 \cdot (\bar{P}_{\pi_O} - \bar{P}_{\pi_{CEP}})/\bar{P}_{\pi_O}$, where $\bar{P}_\pi$ is the average profit of policy $\pi \in \{\pi_{CEP}, \pi_O\}$. Since π_O is unattainable in practice, $\mathcal{R}$ serves as an upper bound on the relative regret with the best achievable performance.

The simulation procedure is as follows. For each instance, we first sample a base rate, say λ^*, from the Gamma distribution $\Lambda \sim \Gamma(\alpha_0, \beta_0)$. Then we compute the expected optimal profit of the Oracle policy using base rate λ^*: $J^*_{\lambda^*}(0, T)$. We then simulate a sample path of an inhomogeneous Poisson process with base rate λ^* that is controlled by the CE policy using the well-known thinning

algorithm of Lewis and Shedler (1979) and obtain the resulting profit. We repeat this 2,000 times to obtain estimates for both $\bar{P}_{\pi_O}$ and $\bar{P}_{\pi_{CEP}}$, respectively, with the width of the 95% confidence intervals being less than 0.1% of the point estimates. We then compute $\bar{\mathscr{R}}$ using $\bar{P}_{\pi_O}$ and $\bar{P}_{\pi_{CEP}}$. Table 3 summarizes the simulation study's main results. This table presents the average and maximum $\bar{\mathscr{R}}$ and the $\bar{P}_{\pi_O}$ across all instances. Detailed results can be found in Appendix D.3.

Table 3 indicates that the CE policy performs excellently with average profit losses of only 0.72% compared to the Oracle policy across a wide range of instances. Additionally, with a maximum profit loss of only 11.63% over all 4050 instances, the CE policy is also robust, providing good performance under varying conditions.

Table 3   Main results certainty-equivalent policy.

Mean $\bar{\mathscr{R}}$	Max $\bar{\mathscr{R}}$	$\bar{P}_{\pi_O}$
0.72	11.63	10.33

It is important to mention that the detailed results in Table 8 in Appendix D.3 indicate that a concave deterioration function and a convex revenue function lead to particularly small profit losses. This is precisely the regime (cf. Corollary 1 (iii)) where optimal condition-based production policies, and thus the CE and Oracle policy, are bang-bang policies (recall that this policy characterization does not depend on the value of λ). Hence, these results suggest that CE policies are robust against unit-to-unit variability in the bang-bang regime; that is, they achieve low relative profit losses in that regime compared to the Oracle policy regardless of the degree of unit-to-unit variability. In the next section, we report on additional numerical analyses that provide strong further evidence for this finding.

6.4. Robustness of Certainty-Equivalent Bang-Bang Policies

Figure 10 contains plots that exemplify behaviour we consistently observed in our numerical experiments. In these plots, we use the base case instance from §5.1.2, except that we compare pairs of (γ, ν) values, fix $\mathbb{E}[\Lambda] = 1$, and vary cv_Λ. We also keep N^{opt} fixed at 0 to isolate and study the effect of the bang-bang structure – and not the effect of re-optimization – on the relative profit loss $\bar{\mathscr{R}}$ (which we compute using the simulation procedure described in the previous subsection). In each plot, one (γ, ν) value corresponds to a bang-bang policy, while the other does not. The plots clearly indicate that $\bar{\mathscr{R}}$ is close to zero when the CE policy is bang-bang (squares) for all values of cv_Λ while it increases in cv_Λ when the CE policy is not bang-bang (circles).

The robustness of a CE bang-bang policy against unit-to-unit variability can be explained as follows. When the Oracle is bang-bang, the CE policy is as well – and vice versa – as this only depends on $f(\cdot)$ and $r(\cdot)$, and not on (the realization of) the base rate λ. Since a bang-bang policy
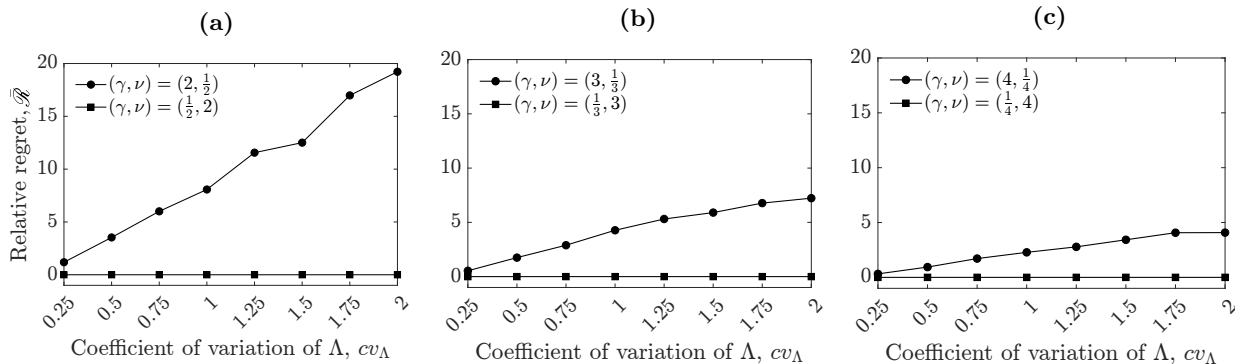
(a)   **(b)**   **(c)**

Figure 10   Comparisons of the relative profit loss (in %) of the CE policy with $N^{\mathrm{opt}} = 0$ as a function of cv_Λ for three pairs of (γ, ν)-values with $f = s^\gamma$ and $r = s^\nu$. In each comparison, the relative profit loss of the CE bang-bang policy is depicted with the solid squares.

can only take two values (0 or $\bar{s}$) in each state (x, t), it is much more likely that the CE and Oracle policy coincide throughout each sample path realization, than when the policy is not bang-bang in which the policy can take on infinitely many values in each state (x, t). As a result, the expected profits of the Oracle and CE policy in the bang-bang regime will also be of similar magnitude. Indeed, we observe in Figure 10 that the relative profit loss of the CE policy is not impacted by unit-to-unit variability in the bang-bang regime.

This finding has an important managerial implication for DMs facing unit-to-unit variability in condition-based production systems. When DMs operate in the regime where a bang-bang policy is optimal (which we established can be verified without knowing the value of λ, see Proposition 5), DMs do not have to take variability in deterioration rates into account, but only require the population's average deterioration rate $\mathbb{E}[\Lambda]$. As DMs in practice can often readily obtain an unbiased estimate of $\mathbb{E}[\Lambda]$ through data, implementing a CE policy with $N^{\mathrm{opt}} = 0$ is not only computationally less demanding and straightforward, it also ensures superior performance regardless of the realization of λ. When the policy is not bang-bang, DMs can rely on a CE policy with $N^{\mathrm{opt}} > 0$ to learn λ and still effectively manage condition-based production systems that have unit-to-unit variability. The numerical results in §6.3 suggest that the CE policy is indeed capable of achieving good overall performance when $N^{\mathrm{opt}} > 0$.

7. Concluding Remarks

Building on advancements in sensor technology and Internet-of-Things connectivity, we addressed two interconnected decision problems for stochastically deteriorating production systems, focusing on real-time monitoring and control. We first studied the operational decision problem of how to adjust the production rate to balance deterioration and revenue accumulation based on the real-time condition signal given that maintenance moments are fixed. We modeled this problem as a continuous-time MDP, and established intuitive monotonic properties of the optimal policy. We

also presented easily verifiable conditions that are sufficient for a bang-bang policy to be optimal. Next, we considered the tactical decision problem of finding the optimal planned maintenance moments given that the optimal condition-based production policy is followed in between. We showed that this decision problem can be easily optimized, thereby enabling the optimal integration of condition-based production on an operational level and maintenance decisions on a tactical level. Finally, we extended our analysis to populations of production systems that exhibit unit-to-unit variability in the sense that the production-deterioration relation is not known and needs to be learned on-the-fly. For the operational decision problem in this setting, we proposed a Bayesian procedure to learn the unknown deterioration rate under any production policy in real time.

Our extensive numerical studies have led to several important insights. Firstly, condition-based production policies enabled by real-time condition monitoring and remote control significantly outperform static policies. Secondly, the integration of tactical maintenance planning with operational condition-based production policies leads to substantial profit increases, as opposed to addressing these two problems sequentially. This underscores the importance of setting maintenance intervals in conjunction with operational condition-based production decisions. Thirdly, in cases where production systems exhibit unit-to-unit variability in the production-deterioration relationship, our Bayesian heuristic policy performs close to an Oracle policy. Notably, our findings highlight a key insight: The heuristic policy demonstrates particularly superior performance in the bang-bang regime, which remains robust even in the presence of high unit-to-unit variability.

References

Alaswad S, Xiang Y (2017) A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety* 157:54–63, ISSN 0951-8320.

Aliprantis CD, Border KC (1994) *Infinite Dimensional Analysis: A Hitchhiker's Guide.* (Springer-Verlag, Berlin).

Araman VF, Caldentey R (2009) Dynamic pricing for nonperishable products with demand learning. *Operations Research* 57(5):1169–1188.

Arts J, Basten R (2018) Design of multi-component periodic maintenance programs with single-component models. *IISE Transactions* 50(7):606–615.

Aydin G, Ziya S (2008) Pricing promotional products under upselling. *Manufacturing & Service Operations Management* 10(3):360–376.

Aydin G, Ziya S (2009) Personalized dynamic pricing of limited inventories. *Operations Research* 57(6):1523–1531.

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

Batun S, Maillart LM (2012) Reassessing tradeoffs inherent to simultaneous maintenance and production planning. *Production and Operations Management* 21(2):396–403.

Bitran GR, Mondschein SV (1997) Periodic pricing of seasonal products in retailing. *Management Science* 43(1):64–79.

Blackwell E, Gambell T, Marya V, Schmitz C (2017) The great re-make: Manufacturing for modern times. Technical report, McKinsey.

Boukas E, Liu Z (2001) Production and maintenance control for manufacturing systems. *IEEE Transactions on Automatic Control* 46(9):1455–1460.

Boukas E, Zhang Q, Yin G (1995) Robust production and maintenance planning in stochastic manufacturing systems. *IEEE Transactions on Automatic Control* 40(6):1098–1102.

Bremaud P (1980) *Point Processes and Queues, Martingale Dynamics* (Springer-Verlag, NY).

Cao P, Li J, Yan H (2012) Optimal dynamic pricing of inventories with stochastic demand and discounted criterion. *European Journal of Operational Research* 217(3):580–588.

Chao X, Chen H, Zheng S (2009) Dynamic capacity expansion for a service firm with capacity deterioration and supply uncertainty. *Operations Research* 57(1):82–93.

Coleman C, Damodaran S, Chandramouli M, Deuel E (2017) Making maintenance smarter: Predictive maintenance and the digital supply network. Technical report, Deloitte.

Conrad C, McClamroch N (1987) The drilling problem: A stochastic modeling and control example in manufacturing. *IEEE Transactions on Automatic Control* 32(11):947–958.

Coopers PW (2014) Sensing the future of the internet of things. Digital IQ snapshot.

De Jonge B, Scarf PA (2020) A review on maintenance optimization. *European Journal of Operational Research* 285(3):805–824.

Derman C (1963) Optimal replacement and maintenance under markovian deterioration with probability bounds on failure. *Management Science* 9(3):478–481.

Dolinšek S, Šuštaršič B, Kopač J (2001) Wear mechanisms of cutting tools in high-speed cutting processes. *Wear* 250(1-12):349–356.

Drent C, Drent M, Arts J, Kapodistria S (2023a) Real-time integrated learning and decision making for cumulative shock degradation. *Manufacturing & Service Operations Management* 25(1):235–253.

Drent C, Drent M, Van Houtum GJ (2023b) Optimal data pooling for shared learning in maintenance operations. *Operations Research Letters* (in press).

Du L, Hu M, Wu J (2022) Sales effort management under all-or-nothing constraint. *Management Science* 68(7):5109–5126.

Elwany AH, Gebraeel NZ, Maillart LM (2011) Structured replacement policies for components with complex degradation processes and dedicated sensors. *Operations Research* 59(3):684–695.

Esary J, Marshall A (1973) Shock models and wear processes. *The Annals of Probability* 1(4):627–649.

Farias VF, Van Roy B (2010) Dynamic pricing with a prior on market response. *Operations Research* 58(1):16–29.

Feng Y, Qiu Y, Crabtree CJ, Long H, Tavner PJ (2013) Monitoring wind turbine gearboxes. *Wind Energy* 16(5):728–740.

Francie KA, Jean-Pierre K, Pierre D, Victor S, Vladimir P (2014) Stochastic optimal control of manufacturing systems under production-dependent failure rates. *International Journal of Production Economics* 150:174–187.

Gallego G, Topaloglu H (2019) *Revenue Management and Pricing Analytics*, volume 209 (New York: Springer).

Gallego G, Van Ryzin G (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science* 40(8):999–1020.

Gallego G, Van Ryzin G (1997) A multiproduct dynamic pricing problem and its applications to network yield management. *Operations Research* 45(1):24–41.

Gebraeel NZ, Lawley MA, Li R, Ryan JK (2005) Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions* 37(6):543–557.

Hao L, Liu K, Gebraeel NZ, Shi J (2015) Controlling the residual life distribution of parallel unit systems through workload adjustment. *IEEE Transactions on Automation Science and Engineering* 14(2):1042–1052.

Hu J, Xiang D (1995) Monotonicity of optimal flow control for failure-prone production systems. *Journal of Optimization Theory and Applications* 86(1):57–71.

Hu S, Hu X, Ye Q (2017) Optimal rebate strategies under dynamic pricing. *Operations Research* 65(6):1546–1561.

Kazaz B, Sloan TW (2008) Production policies under deteriorating process conditions. *IIE Transactions* 40(3):187–205.

Kazaz B, Sloan TW (2013) The impact of process deterioration on production and maintenance policies. *European Journal of Operational Research* 227(1):88–100.

Kim MJ, Makis V (2013) Joint optimization of sampling and control of partially observable failing systems. *Operations Research* 61(3):777–790.

Kleywegt AJ, Papastavrou JD (2001) The dynamic and stochastic knapsack problem with random sized items. *Operations Research* 49(1):26–41.

Lamghari-Idrissi D (2021) *A New After-Sales Service Measure for Stable Customer Operations*. Ph.D. thesis, Eindhoven University of Technology.

Lewis PW, Shedler GS (1979) Simulation of nonhomogeneous Poisson processes by thinning. *Naval Research Logistics Quarterly* 26(3):403–413.

Li C, Tomlin B (2022) After-sales service contracting: Condition monitoring and data ownership. *Manufacturing & Service Operations Management* 24(3):1494–1510.

Liberopoulos G, Caramanis M (1994) Production control of manufacturing systems with production rate-dependent failure rates. *IEEE Transactions on Automatic Control* 39(4):889–895.

Macquart T, Maheri A (2019) A stall-regulated wind turbine design to reduce fatigue. *Renewable Energy* 133:964–970.

Martinelli F (2010) Manufacturing systems with a production dependent failure rate: structure of optimality. *IEEE Transactions on Automatic Control* 55(10):2401–2406.

Mula J, Poler R, García-Sabater J, Lario F (2006) Models for production planning under uncertainty: A review. *International Journal of Production Economics* 103(1):271–285, ISSN 0925-5273.

Nourelfath M, Yalaoui F (2012) Integrated load distribution and production planning in series-parallel multi-state systems with failure rate depending on load. *Reliability Engineering & System Safety* 106:138–145.

Olsen TL, Tomlin B (2020) Industry 4.0: Opportunities and challenges for operations management. *Manufacturing & Service Operations Management* 22(1):113–122.

Rishel R (1989) Controlled wear processes: Modeling optimal control. *Proceedings of the 28th IEEE Conference on Decision and Control*, 724–726 (IEEE).

Sethi SP, Yan H, Zhang H, Zhang Q (2002) Optimal and hierarchical controls in dynamic stochastic manufacturing systems: A survey. *Manufacturing & Service Operations Management* 4(2):133–170.

Shaked M, Shanthikumar JG (2007) *Stochastic Orders* (Springer Science & Business Media).

Sloan TW, Shanthikumar JG (2000) Combined production and maintenance scheduling for a multiple-product, single-machine production system. *Production and Operations Management* 9(4):379–399.

Sloan TW, Shanthikumar JG (2002) Using in-line equipment condition and yield information for maintenance scheduling and dispatching in semiconductor wafer fabs. *IIE Transactions* 34(2):191–209.

Sobczyk K (1987) Stochastic models for fatigue damage of materials. *Advances in Applied Probability* 19(3):652–673.

Spinato F, Tavner PJ, Van Bussel GJ, Koutoulakos E (2009) Reliability of wind turbine subassemblies. *IET Renewable Power Generation* 3(4):387–401.

Sun Q, Ye ZS, Revie M, Walls L (2019) Reliability modeling of infrastructure load-sharing systems with workload adjustment. *IEEE Transactions on Reliability* 68(4):1283–1295.

Tian F, Sun P, Duenyas I (2021) Optimal contract for machine repair and maintenance. *Operations Research* 69(3):916–949.

Topkis DM (1998) *Supermodularity and Complementarity* (Princeton University Press).

Uit het Broek MA, Teunter RH, De Jonge B, Veldman J (2021) Joint condition-based maintenance and condition-based production optimization. *Reliability Engineering & System Safety* 214:107743.

Uit het Broek MA, Teunter RH, De Jonge B, Veldman J, Van Foreest ND (2020) Condition-based production planning: Adjusting production rates to balance output and failure risk. *Manufacturing & Service Operations Management* 22(4):792–811.

Van Dijkhuizen G, van Harten A (1997) Optimal clustering of frequency-constrained maintenance jobs with shared set-ups. *European Journal of Operational Research* 99(3):552–564.

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

Zhang H, Zhang W (2023) Analytical solution to a partially observable machine maintenance problem with obvious failures. *Management Science* 69(7):3993–4015.

Zhang T, Dwight R, El-Akruti K (2015) Condition based maintenance and operation of wind turbines. *Engineering Asset Management-Systems, Professional Practices and Certification*, 1013–1025 (Springer).

Zied H, Sofiene D, Nidhal R (2011) Optimal integrated maintenance/production policy for randomly failing systems with variable failure rate. *International Journal of Production Research* 49(19):5695–5712.

Appendix A: Proofs of Statements

A.1. Proof of Lemma 1

Proof. We use a standard argument in which we consider the dynamics of the deterioration under a chosen setting – essentially an inhomogeneous Poisson process – over a very small time interval.

Consider the dynamics over a small interval of length δt (that is, when the time until the next maintenance moment changes from t to $t - \delta t$) when $t > 0$. Let $x < \xi$. If the decision maker selects production rate $s \in \mathscr{S}$, then with probability $\delta t \lambda f(s) + o(\delta t)$, x changes to $x + 1$, and with probability $1 - \delta t \lambda f(s) + o(\delta t)$, x stays the same. Here, $o(\delta t)$ is a quantity that goes to zero faster than δt. Secondly, if $x = \xi$, then the only admissible action is $s = 0$, so that with probability 1, x stays ξ. In all cases, t changes to $t - \delta t$. By the principle of optimality, for $t > 0$, we have

$$J_\lambda^*(x,t) = \max_{s \in \mathscr{S}} \left[r(s)\delta t + \delta t \lambda f(s) J_\lambda^*(x+1, t-\delta t) + \left(1 - \delta t \lambda f(s)\right) J_\lambda^*(x, t-\delta t) + o(\delta t) \right], \quad (6)$$

when $x < \xi$ and

$$J_\lambda^*(\xi, t) = J_\lambda^*(\xi, t - \delta t), \quad (7)$$

when $x = \xi$. Rearranging (6) and (7), dividing by δt, and taking the limit $\delta t \to 0$, leads to the desired result. The boundary conditions follow from the maintenance cost that is incurred when the system has deterioration level x at the scheduled maintenance moment. □

A.2. Proof of Lemma 2

Proof. We prove each assertion of the statement using a sample path argument. For the monotonicity in x, consider a sample path where the system with deterioration level x is governed by the optimal condition-based production policy for the system with deterioration level $x + 1$ until either the system with deterioration level $x + 1$ breaks down, or until the scheduled maintenance moment, whichever happens first. Note that since the policies are coupled, they will face the same deterioration sample path. If the system with deterioration level $x + 1$ breaks down at time τ before the scheduled maintenance moment, then the system with deterioration level x can still generate revenue over $(T - \tau, 0]$, possibly even with less maintenance costs if the system does not break down. On the other hand, if the scheduled maintenance moment arrives first, the two systems will have generated the same revenue and will have the same maintenance cost. Consequently, the system with deterioration level x generates at least as much revenue and incurs at most the same maintenance costs. Hence, for their expected profits, we have that $J_\lambda^*(x,t) \geq J_\lambda^*(x+1,t)$. Using similar arguments one can also prove the monotonicity in t and λ. □

A.3. Proof of Proposition 1

Proof. We first introduce some additional notation. Let $\Delta_\lambda^2(x,t) \triangleq \Delta_\lambda(x,t) - \Delta_\lambda(x+1,t)$ for $x \leq \xi - 2$. To establish the concavity of $J_\lambda^*(x,t)$ in x for all $t \geq 0$ and each base rate, we need to show that $\Delta_\lambda^2(x,t) \leq 0$ for all $x \leq \xi - 2$ and $t, \lambda \geq 0$.

For $t = 0$, it is easy to verify that $\Delta_\lambda^2(x,0) \leq 0$ for all $x \leq \xi - 2$ and $\lambda \geq 0$, because $-c_m(x)$ is concave in x by assumption.

We now consider $t > 0$. We first construct some useful inequalities that we will use in an inductive proof to establish the result. Let $x \leq \xi - 2$, using (2), we have

$$\frac{\partial J_\lambda^*(x,t)}{\partial t} = r\big(s_\lambda^*(x,t)\big) - \lambda f\big(s_\lambda^*(x,t)\big)\Delta_\lambda(x,t), \tag{8}$$

and since $s_\lambda^*(x,t)$ is also admissible, but not necessarily optimal, for state $(x+1,t)$ with base rate λ, we have

$$\frac{\partial J_\lambda^*(x+1,t)}{\partial t} \geq r\big(s_\lambda^*(x,t)\big) - \lambda f\big(s_\lambda^*(x,t)\big)\Delta_\lambda(x+1,t). \tag{9}$$

Then, subtracting (9) from (8) leads to

$$\frac{\partial \Delta_\lambda(x,t)}{\partial t} \leq -\lambda f\big(s_\lambda^*(x,t)\big)\Delta_\lambda^2(x,t). \tag{10}$$

Similarly, we have

$$\frac{\partial J_\lambda^*(x+2,t)}{\partial t} = r\big(s_\lambda^*(x+2,t)\big) - \lambda f\big(s_\lambda^*(x+2,t)\big)\Delta_\lambda(x+2,t), \tag{11}$$

and since $s_\lambda^*(x+2,t)$ is also admissible, but not necessarily optimal, for state $(x+1,t)$ with base rate λ, we have

$$\frac{\partial J_\lambda^*(x+1,t)}{\partial t} \geq r\big(s_\lambda^*(x+2,t)\big) - \lambda f\big(s_\lambda^*(x+2,t)\big)\Delta_\lambda(x+1,t). \tag{12}$$

Then, subtracting (11) from (12) gives

$$\frac{\partial \Delta_\lambda(x+1,t)}{\partial t} \geq -\lambda f\big(s_\lambda^*(x+2,t)\big)\Delta_\lambda^2(x+1,t). \tag{13}$$

Next, subtracting (13) from (10) leads to

$$\frac{\partial \Delta_\lambda^2(x,t)}{\partial t} \leq -\lambda f\big(s_\lambda^*(x,t)\big)\Delta_\lambda^2(x,t) + \lambda f\big(s_\lambda^*(x+2,t)\big)\Delta_\lambda^2(x+1,t). \tag{14}$$

We now proceed with our inductive proof. We first prove the base case, i.e. $\Delta_\lambda^2(\xi-2,t) \leq 0$ for all $t > 0$ and $\lambda \geq 0$. By (10) we have that

$$\frac{\partial \Delta_\lambda(\xi-2,t)}{\partial t} \leq -\lambda f\big(s_\lambda^*(\xi-2,t)\big)\Delta_\lambda^2(\xi-2,t). \tag{15}$$

Note that due to the linearity of the derivative, we have

$$\frac{\partial \Delta_\lambda(\xi-1,t)}{\partial t} = \frac{\partial J_\lambda^*(\xi-1,t)}{\partial t} - \frac{\partial J_\lambda^*(\xi,t)}{\partial t} = \frac{\partial J_\lambda^*(\xi-1,t)}{\partial t} \geq 0, \tag{16}$$

where the second equality follows from (2) for $x = \xi$ and the inequality follows from Lemma 2. Equation (15) minus Equation (16) on both sides leads to

$$\frac{\partial \Delta_\lambda^2(\xi-2,t)}{\partial t} \leq -\lambda f\big(s_\lambda^*(\xi-2,t)\big)\Delta_\lambda^2(\xi-2,t). \tag{17}$$

Applying Grönwall's Lemma to Inequality (17) leads to

$$\Delta_\lambda^2(\xi-2,t) \leq \Delta_\lambda^2(\xi-2,0) \cdot \exp\left(\int_0^t -\lambda f(s_\lambda^*(\xi-2,u))du\right). \tag{18}$$

Since $\Delta_\lambda^2(\xi-2,0) \leq 0$, (recall that $-c_m(x)$ is concave in x by assumption), and because $\exp\left(\int_0^t -\lambda f(s_\lambda^*(\xi-2,u))du\right) > 0$ for all $t > 0$ and $\lambda \geq 0$, we have that $\Delta_\lambda^2(\xi-2,t) \leq 0$, which proves the base case.

Assume now inductively that $\Delta_\lambda^2(x,t) \leq 0$ for an $x \leq \xi - 2$. We will show that this implies that $\Delta_\lambda^2(x-1,t) \leq 0$ for all $t > 0$ and all $\lambda \geq 0$.

For $(x-1,t)$, we have

$$\frac{\partial \Delta_\lambda^2(x-1,t)}{\partial t} \leq -\lambda f\big(s_\lambda^*(x-1,t)\big)\Delta_\lambda^2(x-1,t) + \lambda f\big(s_\lambda^*(x+1,t)\big)\Delta_\lambda^2(x,t)$$
$$\leq -\lambda f\big(s_\lambda^*(x-1,t)\big)\Delta_\lambda^2(x-1,t). \tag{19}$$

The first inequality is Equation (14), and the second inequality is due to the induction hypothesis. Applying Grönwall's Lemma to Inequality (19) leads to

$$\Delta_\lambda^2(x-1,t) \leq \Delta_\lambda^2(x-1,0) \cdot \exp\left(\int_0^t -\lambda f(s_\lambda^*(x-1,u))du\right). \tag{20}$$

Since $\Delta_\lambda^2(x-1,0) \leq 0$ for all $x \leq \xi - 2$ (again by the concavity assumption of $-c_m(x)$ in x) and because $\exp\left(\int_0^t -\lambda f(s_\lambda^*(x-1,u))du\right) > 0$ for all $t > 0$ and $\lambda \geq 0$, we have that $\Delta_\lambda^2(x-1,t) \leq 0$ for all $t > 0$ and all $\lambda \geq 0$. $\qquad\square$

A.4. Proof of Proposition 2

Proof. In this proof, we employ a commonly used transformation of the continuous model into an analog discrete dynamic program (see e.g., Bitran and Mondschein 1997, Aydin and Ziya 2008, Hu et al. 2017). We are allowed to do so because it is known that this discrete dynamic program is guaranteed to converge uniformly to the continuous model (cf. Kleywegt and Papastavrou 2001).

To this end, we divide the entire planning horizon into $\hat{T} = \frac{T}{\delta t}$ periods, each of which is short enough that the probability of more than one deterioration increment during an interval of length

δt is negligible. Indeed, note that since the function $f(\cdot)$ is bounded, we have that for each choice of production rate s, the probability that more than 1 arrival will occur in a small interval of length δt is of order $o(\delta t)$.

Let $J_\lambda^*(x, k)$ denote the maximum expected revenue when the deterioration level is x and there are k periods to go until the planned maintenance moment, and the base rate is λ. We then need to show that the following inequality holds

$$J_\lambda^*(x-1, k) - J_\lambda^*(x, k) \geq J_\lambda^*(x-1, k-1) - J_\lambda^*(x, k-1). \tag{21}$$

The dynamic program for the revenue maximization problem over the $\hat{T}$-period horizon is given by

$$J_\lambda^*(x, k) = \max_{s \in \mathscr{S}} \left[r(s)\delta t + \lambda f(s)\delta t J_\lambda^*(x+1, k-1) + \big(1 - \lambda f(s)\delta t\big) J_\lambda^*(x, k-1) \right], \tag{22}$$

with boundary conditions $J_\lambda^*(x, 0) = -c_m(x)$ for all $x \in \mathscr{X}$.

Let $s_\lambda^*(x, k)$ denote the optimal production rate for state (x, k) when the base rate is λ. We then have, using (22),

$$J_\lambda^*(x, k) = r\big(s_\lambda^*(x, k)\big)\delta t + \lambda f\big(s_\lambda^*(x, k)\big)\delta t J_\lambda^*(x+1, k-1) + \big(1 - \lambda f\big(s_\lambda^*(x, k)\big)\delta t\big) J_\lambda^*(x, k-1). \tag{23}$$

Subtracting $J_\lambda^*(x, k-1)$ from both sides of (23) leads to the following equality

$$J_\lambda^*(x, k) - J_\lambda^*(x, k-1) = r\big(s_\lambda^*(x, k)\big)\delta t + \lambda f\big(s_\lambda^*(x, k)\big)\delta t \big(J_\lambda^*(x+1, k-1) - J_\lambda^*(x, k-1)\big). \tag{24}$$

Since $s_\lambda^*(x, k)$ is admissible in state $(x-1, k)$ when the base rate is λ but not necessarily optimal, we obtain the following inequality

$$J_\lambda^*(x-1, k) \geq$$
$$r\big(s_\lambda^*(x, k)\big)\delta t + \lambda f\big(s_\lambda^*(x, k)\big)\delta t J_\lambda^*(x, k-1) + \big(1 - \lambda f\big(s_\lambda^*(x, k)\big)\delta t\big) J_\lambda^*(x-1, k-1). \tag{25}$$

Subtracting $J_\lambda^*(x-1, k-1)$ from both sides of (25) and then using the concavity of $J_\lambda^*(x, k)$ in x (see Proposition 1) yields

$$J_\lambda^*(x-1, k) - J_\lambda^*(x-1, k-1)$$
$$\geq r\big(s_\lambda^*(x, k)\big)\delta t + \lambda f\big(s_\lambda^*(x, k)\big)\delta t \big(J_\lambda^*(x, k-1) - J_\lambda^*(x-1, k-1)\big)$$
$$\geq r\big(s_\lambda^*(x, k)\big)\delta t + \lambda f\big(s_\lambda^*(x, k)\big)\delta t \big(J_\lambda^*(x+1, k-1) - J_\lambda^*(x, k-1)\big). \tag{26}$$

Observe that the right hand side of (26) can be replaced by (24), which yields the following inequality

$$J_\lambda^*(x-1, k) - J_\lambda^*(x-1, k-1) \geq J_\lambda^*(x, k) - J_\lambda^*(x, k-1),$$

which, after rearranging terms, equals (21). $\qquad\square$

A.5. Proof of Proposition 3

Proof. We need to show that $\frac{\partial J_\lambda^*(x,t)}{\partial t}$ is decreasing in t. Let $u > t$. We have

$$\frac{\partial J_\lambda^*(x,u)}{\partial u} = r\big(s_\lambda^*(x,u)\big) - \lambda f\big(s_\lambda^*(x,u)\big)\Delta_\lambda(x,u), \tag{27}$$

and since $s_\lambda^*(x,u)$ is also admissible, but not necessarily optimal (recall that $s_\lambda^*(x,u)$ is the optimal production rate for state (x,u) when the base rate is λ), for state (x,t) when the base rate is λ, we have

$$\frac{\partial J_\lambda^*(x,t)}{\partial t} \geq r\big(s_\lambda^*(x,u)\big) - \lambda f\big(s_\lambda^*(x,u)\big)\Delta_\lambda(x,t). \tag{28}$$

Subtracting (28) from (27) yields

$$\frac{\partial J_\lambda^*(x,t)}{\partial t} - \frac{\partial J_\lambda^*(x,u)}{\partial u} \geq \lambda f\big(s_\lambda^*(x,u)\big)\Big(\Delta_\lambda(x,u) - \Delta_\lambda(x,t)\Big) \geq 0,$$

where the last inequality is because $\Delta_\lambda(x,u) \geq \Delta_\lambda(x,t)$ for $u > t$ (see Proposition 2). $\qquad\square$

A.6. Proof of Lemma 4

Proof. Let $x_1 \leq x_2$ and $\theta_1 \leq \theta_2$. Since $Z(\theta_1) \leq_{\mathrm{st}} Z(\theta_2)$, there exist two random variables, $\tilde{Z}(\theta_1)$ and $\tilde{Z}(\theta_2)$, on the same probability space, such that $\tilde{Z}(\theta_1) =_{\mathrm{st}} Z(\theta_1)$, $\tilde{Z}(\theta_2) =_{\mathrm{st}} Z(\theta_2)$, and $\tilde{Z}(\theta_1) \leq \tilde{Z}(\theta_2)$ almost surely (here, $=_{\mathrm{st}}$ denotes equality in law) (see, e.g., Shaked and Shanthikumar 2007, Theorem 1.A.1.). We therefore have:

$$\begin{aligned}
\mathbb{E}[f(x_1 + Z(\theta_1),\theta_1)] - \mathbb{E}[f(x_2 + Z(\theta_1),\theta_1)] &= \mathbb{E}[f(x_1 + \tilde{Z}(\theta_1),\theta_1) - f(x_2 + \tilde{Z}(\theta_1),\theta_1)] \\
&\leq \mathbb{E}[f(x_1 + \tilde{Z}(\theta_1),\theta_2) - f(x_2 + \tilde{Z}(\theta_1),\theta_2)] \\
&\leq \mathbb{E}[f(x_1 + \tilde{Z}(\theta_2),\theta_2) - f(x_2 + \tilde{Z}(\theta_2),\theta_2)] \\
&= \mathbb{E}[f(x_1 + Z(\theta_2),\theta_2)] - \mathbb{E}[f(x_2 + Z(\theta_2),\theta_2)],
\end{aligned}$$

where the first inequality is from the submodularity of $f(x,\theta)$, and the second inequality is from the concavity in x of $f(x,\theta)$. Hence, $\mathbb{E}[f(x + Z(\theta),\theta)]$ is submodular in (x,θ).

The preservation of concavity under expectation is a well-known result. For completeness, we provide a short proof here. Let $x_1, x_2 \in \mathbb{R}$ and $\alpha \in [0,1]$. We then have

$$\begin{aligned}
\mathbb{E}[f(\alpha x_1 + (1-\alpha)x_2 + Z(\theta),\theta)] &= \mathbb{E}[f(\alpha x_1 + (1-\alpha)x_2 + \alpha Z(\theta) + (1-\alpha)Z(\theta),\theta)] \\
&= \mathbb{E}[f(\alpha\big(x_1 + Z(\theta)\big) + (1-\alpha)\big(x_2 + Z(\theta)\big),\theta)] \\
&\geq \mathbb{E}[\alpha f\big(x_1 + Z(\theta),\theta\big) + (1-\alpha)f\big(x_2 + Z(\theta),\theta\big)] \\
&= \alpha\mathbb{E}[f\big(x_1 + Z(\theta),\theta\big)] + (1-\alpha)\mathbb{E}[f\big(x_2 + Z(\theta),\theta\big)],
\end{aligned}$$

where the inequality follows from the concavity of $f(x,\theta)$ in x. $\qquad\square$

A.7. Proof of Lemma 3

Proof.   We need to show that for $x_1 \leq x_2$ and $y_1 \leq y_2$ the following inequality holds:

$$g(y_1)f(x_1, y_1) - g(y_2)f(x_1, y_2) \leq g(y_1)f(x_2, y_1) - g(y_2)f(x_2, y_2).$$

To this end, let $x_1 \leq x_2$ and $y_1 \leq y_2$, we then have

$$g(y_1)f(x_1, y_1) - g(y_2)f(x_1, y_2) - g(y_1)f(x_2, y_1) + g(y_2)f(x_2, y_2)$$
$$= g(y_1)\big(f(x_1, y_1) - f(x_2, y_1)\big) - g(y_2)\big(f(x_1, y_2) - f(x_2, y_2)\big)$$
$$\leq g(y_2)\big(f(x_1, y_1) - f(x_2, y_1) - f(x_1, y_2) + f(x_2, y_2)\big) \leq 0.$$

The first inequality is because $g(y_1) \leq g(y_2)$ and $f(x_1, y_1) \geq f(x_2, y_1)$, and the second inequality is due to the submodularity of $f(x, y)$ in (x, y). $\square$

A.8. Proof of Proposition 4

Proof.   Again, we use the analog discrete dynamic program for the revenue maximization problem over the $\hat{T}$-period horizon and rewrite it in expectation form:

$$J_\lambda^*(x, k) = \max_{s \in \mathscr{S}} \Big[r(s)\delta t + \mathbb{E}[J_\lambda^*(x + Z(\lambda, s), k - 1)] \Big], \tag{29}$$

with boundary conditions $J_\lambda^*(x, 0) = -c_m(x)$ for all $x \in \mathscr{X}$, and where $Z(\lambda, s)$ is a Bernoulli random variable with success probability $p = \delta t \lambda f(s)$.

We need to show that for all $k \geq 0$, $J_\lambda^*(x, k)$ is submodular in (x, λ). We prove this by induction on k. Note that for the base case $k = 0$, $J_\lambda^*(x, 0) = -c_m(x)$ is evidently submodular in (x, λ). Assume that for some $k - 1$, $J_\lambda(x, k - 1)$ is submodular in (x, λ), we will now show that $J_\lambda(x, k)$ is submodular in (x, λ).

Let $g_\lambda(s, x, k) \triangleq r(s)\delta t + \mathbb{E}[J_\lambda^*(x + Z(\lambda, s), k - 1)]$. Since $Z(\lambda, s)$ is stochastically increasing in λ, the induction hypothesis and the fact that $J_\lambda^*(x, k - 1)$ is concave in x (see Proposition 1), we can employ Lemma 4 to conclude that $g_\lambda(s, x, k)$ is submodular in (x, λ) for all s.

Now, let $s_\lambda^*(x, k) \in \arg\max_{s \in \mathscr{S}}\{g_\lambda(s, x, k)\}$, and let $x_1 < x_2$, and $\lambda_1 < \lambda_2$, we then have:

$$J_{\lambda_1}^*(x_1, k) + J_{\lambda_2}^*(x_2, k) = g_{\lambda_1}\big(s_{\lambda_1}^*(x_1, k), x_1, k\big) + g_{\lambda_2}\big(s_{\lambda_2}^*(x_2, k), x_2, k\big)$$
$$\leq g_{\lambda_2}\big(s_{\lambda_1}^*(x_1, k), x_1, k\big) + g_{\lambda_1}\big(s_{\lambda_2}^*(x_2, k), x_2, k\big)$$
$$\leq g_{\lambda_2}\big(s_{\lambda_2}^*(x_1, k), x_1, k\big) + g_{\lambda_1}\big(s_{\lambda_1}^*(x_2, k), x_2, k\big)$$
$$= J_{\lambda_2}^*(x_1, k) + J_{\lambda_1}^*(x_2, k),$$

where the first inequality is due to the submodularity of $g_\lambda(s, x, k)$ in (x, λ) for all s, and the second inequality is due to the suboptimality of $s_{\lambda_1}^*(x_1, k)$ for (x_1, λ_2) and $s_{\lambda_2}^*(x_2, k)$ for (x_2, λ_1), respectively. Hence $J_\lambda^*(x, k)$ is submodular in (x, λ). $\square$

A.9. Proof of Theorem 1

Proof. The proof is divided into three parts, focused on the monotonicity in x, t, and λ, respectively. In each part we use the notation $g_\lambda(s, x, t) = r(s) - \lambda f(s)\Delta_\lambda(x, t)$, that we previously introduced.

Proof of monotonicity in x: For notational clarity we drop the subscript λ. Let $\hat{s} \triangleq s^*(x, t)$ and consider $s > \hat{s}$, we have that

$$
\begin{aligned}
&g(\hat{s}, x+1, t) - g(\hat{s}, x, t) + g(s, x, t) - g(s, x+1, t) \\
&= -\lambda f(\hat{s})\Delta(x+1, t) + \lambda f(\hat{s})\Delta(x, t) - \lambda f(s)\Delta(x, t) + \lambda f(s)\Delta(x+1, t) \\
&= \lambda f(\hat{s})\Delta^2(x, t) - \lambda f(s)\Delta^2(x, t) \\
&= \lambda\big(f(\hat{s}) - f(s)\big)\Delta^2(x, t) \geq 0.
\end{aligned}
\tag{30}
$$

The inequality follows from the increasing property of $f(\cdot)$ with $s > \hat{s}$, and since $\Delta^2(x, t) \leq 0$ (see Proposition 1). Rearranging Inequality (30) leads to

$$
g(\hat{s}, x+1, t) - g(s, x+1, t) \geq g(\hat{s}, x, t) - g(s, x, t) \geq 0,
\tag{31}
$$

where the second inequality is because $\hat{s} \triangleq s^*(x, t)$. We can therefore conclude from (31), that any $s > \hat{s}$ cannot be optimal for $(x+1, t)$, hence $s^*(x+1, t) \leq s^*(x, t)$.

Proof of monotonicity in t: For notational clarity we again drop the subscript λ. Let $u < t$ and $\hat{s} \triangleq s^*(x, u)$, and consider $s > \hat{s}$. We have that

$$
\begin{aligned}
&g(\hat{s}, x, t) - g(\hat{s}, x, u) + g(s, x, u) - g(s, x, t) \\
&= -\lambda f(\hat{s})\Delta(x, t) + \lambda f(\hat{s})\Delta(x, u) - \lambda f(s)\Delta(x, u) + \lambda f(s)\Delta(x, t) \\
&= \lambda\big(f(\hat{s}) - f(s)\big)\big(\Delta(x, u) - \Delta(x, t)\big) \geq 0.
\end{aligned}
\tag{32}
$$

The inequality follows from the increasing property of $f(\cdot)$ and $s > \hat{s}$, and since $\Delta(x, u) \leq \Delta(x, t)$ for $u < t$ (see Proposition 2). Rearranging Inequality (32) leads to

$$
g(\hat{s}, x, t) - g(s, x, t) \geq g(\hat{s}, x, u) - g(s, x, u) \geq 0,
\tag{33}
$$

where the second inequality is because $\hat{s} \triangleq s^*(x, u)$. We can therefore conclude from (33), that any $s > \hat{s}$ cannot be optimal for (x, t), hence $s^*(x, t) \leq s^*(x, u)$ for $u < t$.

Proof of monotonicity in λ: In this part, we use the analog discrete dynamic program. In light of the well-known Topkis' Theorem (see, for instance, Topkis 1998, Theorem 2.8.1), we need to show that $g_\lambda(s, x, k) = r(s)\delta t + \lambda f(s)\delta t J_\lambda^*(x+1, k-1) + (1 - \lambda f(s)\delta t)J_\lambda^*(x, k-1)$ is submodular in (s, λ) to conclude that $s_\lambda^*(x, k)$ is decreasing in λ.

Let $s_1 \le s_2$ and $\lambda_1 \le \lambda_2$. We then need to show that

$$g_{\lambda_1}(s_1, x, k) - g_{\lambda_2}(s_1, x, k) - g_{\lambda_1}(s_2, x, k) + g_{\lambda_2}(s_2, x, k) \le 0.$$

We have (after some algebraic manipulations):

$$g_{\lambda_1}(s_1, x, k) - g_{\lambda_2}(s_1, x, k) - g_{\lambda_1}(s_2, x, k) + g_{\lambda_2}(s_2, x, k)$$
$$= \lambda_1 f(s_1) J^*_{\lambda_1}(x+1, k-1) - \lambda_1 f(s_1) J^*_{\lambda_1}(x, k-1)$$
$$- \lambda_2 f(s_1) J^*_{\lambda_2}(x+1, k-1) + \lambda_2 f(s_1) J^*_{\lambda_2}(x, k-1)$$
$$- \lambda_1 f(s_2) J^*_{\lambda_1}(x+1, k-1) + \lambda_1 f(s_2) J^*_{\lambda_1}(x, k-1)$$
$$+ \lambda_2 f(s_2) J^*_{\lambda_2}(x+1, k-1) - \lambda_2 f(s_2) J^*_{\lambda_2}(x, k-1)$$
$$= \big(f(s_1) - f(s_2)\big)\big(\lambda_1 J^*_{\lambda_1}(x+1, k-1) - \lambda_1 J^*_{\lambda_1}(x, k-1) - \lambda_2 J^*_{\lambda_2}(x+1, k-1) + \lambda_2 J^*_{\lambda_2}(x, k-1)\big).$$

Since $f(s_1) - f(s_2) \le 0$ ($f(\cdot)$ is increasing and $s_2 \ge s_1$), $g_\lambda(s, x, k)$ is submodular in (s, λ) if

$$\lambda_1 J^*_{\lambda_1}(x+1, k-1) - \lambda_1 J^*_{\lambda_1}(x, k-1) - \lambda_2 J^*_{\lambda_2}(x+1, k-1) + \lambda_2 J^*_{\lambda_2}(x, k-1) \ge 0,$$

that is, if $\lambda J^*_\lambda(x, k-1)$ is submodular in (x, λ).

Note that $J^*_\lambda(x, k-1)$ is submodular in (x, λ) (Proposition 4) and decreasing in x (Lemma 2), hence by applying Lemma 3 with $g(\lambda) \triangleq \lambda$ and $f(x, \lambda) \triangleq J^*_\lambda(x, k-1)$, we can conclude that $\lambda J^*_\lambda(x, k-1)$ is submodular in (x, λ).

$\square$

A.10. Proof of Proposition 5

Proof. Recall that for a state (x, t), we are interested in the following optimization problem:

$$\max_{s \in [0, \bar{s}]} \Big[r(s) - \lambda f(s) \Delta_\lambda(x, t) \Big]. \tag{34}$$

Suppose the contrary of the proposition; that is, if $\frac{r(s)}{f(s)} \le \frac{r(\bar{s})}{f(\bar{s})}$ for all $s \in (0, \bar{s})$ then the optimal policy has not a bang-bang structure. Let $s \in (0, \bar{s})$ be the optimal production rate. We will now show that this leads to a contradiction.

Since $s \in (0, \bar{s})$ is optimal, we have by Equation (34):

$$r(s) - \lambda f(s) \Delta_\lambda(x, t) > r(0) - \lambda f(0) \Delta_\lambda(x, t), \text{ and}$$
$$r(s) - \lambda f(s) \Delta_\lambda(x, t) > r(\bar{s}) - \lambda f(\bar{s}) \Delta_\lambda(x, t).$$

These inequalities are equivalent to, respectively,

$$r(s) > r(0) + \lambda \Delta_\lambda(x, t) \left(f(s) - f(0)\right), \text{ and} \tag{35}$$
$$r(s) > r(\bar{s}) + \lambda \Delta_\lambda(x, t) \left(f(s) - f(\bar{s})\right). \tag{36}$$

Multiplying (35) with $\left(1 - \frac{f(s)}{f(\bar{s})}\right)$ and (36) with $\frac{f(s)}{f(\bar{s})}$ on both sides, and adding them together leads to

$$
\begin{aligned}
r(s) &> \left(1 - \frac{f(s)}{f(\bar{s})}\right) \cdot (r(0) + \lambda\Delta_\lambda(x,t)\left(f(s) - f(0)\right)) + \left(\frac{f(s)}{f(\bar{s})}\right) \cdot (r(\bar{s}) + \lambda\Delta_\lambda(x,t)\left(f(s) - f(\bar{s})\right)) \\
&= \left(1 - \frac{f(s)}{f(\bar{s})}\right) \cdot \lambda\Delta_\lambda(x,t)f(s) + \frac{f(s)}{f(\bar{s})}r(\bar{s}) + \left(\frac{f(s)}{f(\bar{s})}\right) \cdot \lambda\Delta_\lambda(x,t)\left(f(s) - f(\bar{s})\right) \\
&= \lambda\Delta_\lambda(x,t)f(s) - \left(\frac{f(s)}{f(\bar{s})}\right) \cdot \lambda\Delta_\lambda(x,t)f(\bar{s}) + \left(\frac{f(s)}{f(\bar{s})}\right) \cdot r(\bar{s}) \\
&= \frac{f(s)}{f(\bar{s})} \cdot r(\bar{s}).
\end{aligned}
\tag{37}
$$

Rewriting Inequality (37) leads to $\frac{r(s)}{f(s)} > \frac{r(\bar{s})}{f(\bar{s})}$, which is a contradiction. Hence, s cannot be optimal.
□

A.11. Proof of Corollary 1

Proof.   In light of Proposition 5, we need to show that the inequality $\frac{r(s)}{f(s)} \leq \frac{r(\bar{s})}{f(\bar{s})}$ holds for all $s \in (0, \bar{s})$ when either condition (i), (ii), or (iii) holds.

It is easy to verify that the above inequality holds when $r(s) = f(s)$ for all $s \in \mathscr{S}$, hence yielding condition (i). For condition (ii), we first note that $f(\cdot)$ and $r(\cdot)$ are assumed to be increasing functions of s (see Assumption 1), and thus $a > 0$ and $b > 0$. We write $\frac{r(s)}{f(s)} = \frac{a}{b}s^{\alpha-\beta}$, which is non-decreasing in s for $ab \geq 0$ and $\alpha \geq \beta$. This implies that $\frac{a}{b}s^{\alpha-\beta} \leq \frac{a}{b}\bar{s}^{\alpha-\beta}$ for all $s \in (0, \bar{s})$ under condition (ii) and Assumption 1. We now focus on condition (iii). For any $0 < s < \bar{s}$, let $s = t \cdot \bar{s}$, so that $0 < t < 1$. The convexity of $r(\cdot)$ yields $t \cdot r(\bar{s}) + (1 - t) \cdot r(0) \geq r(t \cdot \bar{s}) = r(s)$, which after using $r(0) \leq 0$, leads to

$$
t \cdot r(\bar{s}) \geq r(s).
\tag{38}
$$

Likewise, using the concavity of $f(\cdot)$ and $f(0) = 0$, we obtain

$$
t \cdot f(\bar{s}) \leq f(s).
\tag{39}
$$

Combining Inequality (38) and Inequality (39) leads to $\frac{r(\bar{s})}{f(\bar{s})} \geq \frac{r(s)}{f(s)}$ for all $s \in (0, \bar{s})$. Note that this result can also be obtained by applying Bauer's maximum principle (see, e.g., Theorem 7.69 of Aliprantis and Border 1994). When $r(\cdot)$ is convex and $f(\cdot)$ is concave, our objective function $r(s) - \lambda f(s)\Delta_\lambda(x,t)$ is convex, and since $[0, \bar{s}]$ is a compact and convex set, its maximum must be attained at the boundary of the set $[0, \bar{s}]$.   □

A.12. Proof of Theorem 2

Proof. We use the shorthand notation $J(T) \triangleq J_\lambda^*(0, T)$. Taking the first and second derivative of $g(T)$ with respect to T yields

$$g'(T) = \frac{T J'(T) - J(T)}{T^2}, \tag{40}$$

and, after some algebraic simplifications,

$$g''(T) = \frac{T^2 J''(T) - 2 \cdot \left(T J'(T) - J(T)\right)}{T^3}. \tag{41}$$

Now observe that for any t with $g'(t) = 0$, we have that $t J'(t) - J(t) = 0$ (see Equation (40)), and hence that $g''(t) = \frac{t^2 J''(t) - 2 \cdot 0}{t^3} = \frac{J''(t)}{t}$ (see Equation (41)). We further know that $J(t)$ is concave in t, i.e. $J''(t) \leq 0$ for all $t \geq 0$ (see Proposition 3), so that the following condition holds

$$g''(t) \leq 0, \text{ for all } t \geq 0 \text{ such that } g'(t) = 0. \tag{42}$$

We now proceed with the proof. Let T^* be a strict local maximizer of $g(T)$. Let us assume the contrary of the result, i.e. that T^* is not a strict global maximizer. We will show that this leads to a contradiction.

Let T' denote the strict global maximizer and consider the case that $T' > T^*$ (the proof follows verbatim for the case that $T' < T^*$). By construction, $g(T)$ attains it maximum over the closed interval $[T^*, T']$ in T'. By the extreme value theorem, $g(T)$ also attains a minimum over this closed interval $[T^*, T']$ at least once. Let t_0 be the smallest number in $[T^*, T']$ such that $g(t_0)$ is a local minimum. Observe that since T^* is a strict local maximizer, there exists some $\epsilon_0 > 0$ such that $g(T^*) > g(T^* + \epsilon)$ for all $0 < \epsilon < \epsilon_0$, implying that $t_0 \in [T^* + \epsilon, T')$ for some $\epsilon > 0$. As $g(T)$ is differentiable and continuous in $[T^*, T']$, we have $g'(t_0) = 0$. But by (42) this implies that $g''(t_0) \leq 0$, which contradicts that t_0 is a local minimizer. Thus T^* must be a strict global maximizer. $\square$

A.13. Proof of Proposition 6

Proof. Let $f_t(\boldsymbol{s}_t) \triangleq \int_{u=0}^{t} f(s_u) du$, so that the mean of the deterioration process at time t under an employed production policy $\boldsymbol{s}_t$ and given λ is equal to $\lambda f_t(\boldsymbol{s}_t)$ (this follows from the fact that the deterioration process is a non-homogenous Poisson process with rate $\lambda f(s_t)$ at time t). We can then obtain the posterior distribution at time t using Bayes' theorem. That is, the posterior distribution at time t is proportional to the likelihood function, which is a Poisson distribution with mean $\lambda f_t(\boldsymbol{s}_t)$, times the prior distribution:

$$\varphi_u(\lambda | \boldsymbol{s}, y, \alpha_0, \beta_0) \propto \mathbb{P}[Y_u = y | \lambda, \boldsymbol{s}] \cdot \varphi_0(\lambda | \alpha_0, \beta_0) = \frac{(\lambda f_t(\boldsymbol{s}))^y e^{-\lambda f_t(\boldsymbol{s})}}{y!} \frac{\beta_0^{\alpha_0} \lambda^{\alpha_0 - 1} e^{-\beta_0 \lambda}}{\Gamma(\alpha_0)}$$

$$= \frac{\left(f_t(\boldsymbol{s})\right)^y}{y!} \frac{\beta_0^{\alpha_0} \lambda^{y + \alpha_0 - 1} e^{-\lambda(\beta_0 + f_t(\boldsymbol{s}))}}{\Gamma(\alpha_0)} \propto \lambda^{\alpha_0 + y - 1} e^{-\lambda(\beta_0 + f_t(\boldsymbol{s}))}.$$

Note that this is again a Gamma distribution with parameters $\alpha_0 + y$ and $\beta_0 + f_t(\boldsymbol{s}_t)$. $\square$

Appendix B: Extension to Multiple Production Systems

In Remark 2, we explained that by appropriately choosing the revenue function r, one can model situations where a DM aims to satisfy a (continuous) demand rate D with production rate s (by penalizing the downfall $(D-s)^+$ and rewarding the surplus $(s-D)^+$). One natural extension to this case is when this demand rate D can be satisfied by multiple production systems together. In this section, we discuss this extension.

B.1. Problem Formulation

We have N production systems. Each individual system $i \in \{1, 2, \ldots, N\}$ has base rate λ_i and deterioration function f_i. For ease of exposition we assume that all systems have the same production rate set $\mathscr{S}$, failure threshold ξ, and maintenance cost function c_m. (These can also depend on i leading only to additional notation). Let $\boldsymbol{s} \triangleq (s_1, s_2, \ldots, s_n) \in \mathscr{S}^N$ denote a vector (bold notation indicates vectors) of production rates, where s_i is the production rate of system i. As the DM aims to achieve the common demand rate D with the cumulative production rate from all systems, we let the revenue function $r : \mathscr{S}^N \to \mathbb{R}_+$ be of the form:

$$r(\boldsymbol{s}) \triangleq b \cdot \left(\sum_{i=1}^{N} s_i - D \right)^+ - p \cdot \left(D - \sum_{i=1}^{N} s_i \right)^+ , \tag{43}$$

where p is a penalty cost rate per unit of shortfall, while b is a bonus rate per unit of surplus (we discuss this in Remark 2 for the single system case). In this section, we model the setting that the bonus (penalty) is linear in the surplus (shortfall), but the revenue function easily generalizes to any increasing, continuous function. Since we have a common demand rate D, but differing production systems in terms of their base rates and deterioration functions, the DM is interested in determining how to allocate production among these various systems in a way that best meets demand during the planning horizon while also minimizing maintenance costs incurred at the end of the planning horizon.

Let $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_N) \in \mathscr{X}^N$ be the vector of deterioration levels for all systems, where x_i is the deterioration level of system i. Again, we count time backwards, so that at the start of the planning horizon $t = T$, while at the planned maintenance moment we have $t = 0$. The feasible set of production decisions at any time $t < T$ given $\boldsymbol{x}$ is denoted by $\mathscr{S}(\boldsymbol{x}) \triangleq \{\boldsymbol{s} \in \mathscr{S}^N : s_i = 0 \text{ if } x_i = \xi\}$, which means that if system i is failed, i.e. $x_i = \xi$, we can only turn it off. For any time to planned-maintenance t, let $J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}, t)$ be the optimal expected profit when the systems have deterioration levels $\boldsymbol{x}$ and base rates $\boldsymbol{\lambda} \triangleq (\lambda_1, \lambda_2, \ldots, \lambda_N)$. We shall now derive the HJB equation for $J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}, t)$ by following a similar argument as in the proof of Lemma 1.

Assume that the state is $(\boldsymbol{x}, t) \in \mathscr{X}^N \times [T, 0)$ and we choose production rates $\boldsymbol{s} \in \mathscr{S}(\boldsymbol{x})$. The deterioration process is now an N-dimensional non-homogeneous Poisson process with rate vector

$(\lambda_i f_i(s_i))_{i \in \{1,2,\ldots,N\}}$, so we can consider a small interval of length δt such that the probability of a deterioration increment at system i is approximately $\lambda_i f_i(s_i)\delta t$. We then have by the principle of optimality:

$$
J_{\boldsymbol{\lambda}}^*(\boldsymbol{x},t) = \max_{\boldsymbol{s}\in\mathscr{S}(\boldsymbol{x})} \left[r(\boldsymbol{s})\delta t + \sum_{i=1}^{N} \underbrace{\delta t \lambda_i f_i(s_i) J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}+\boldsymbol{e}_i, t-\delta t)}_{\text{system } i \text{ deteriorates}} + \overbrace{\left(1 - \sum_{i=1}^{N} \delta t \lambda_i f_i(s_i)\right) J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}, t-\delta t)}^{\text{no deterioration}} + o(\delta t) \right],
$$
(44)

where $o(\delta t)$ is a quantity that goes to zero faster than δt, and where $\boldsymbol{e}_i$ denotes the i-th standard basis vector in $\mathbb{R}^N$. Subtracting $J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}, t-\delta t)$ from both sides of (44), dividing by δt and taking the limit $\delta t \to 0$ leads to:

$$
\frac{\delta J_{\boldsymbol{\lambda}}^*(\boldsymbol{x},t)}{\delta t} = \max_{\boldsymbol{s}\in\mathscr{S}(\boldsymbol{x})} \left[r(\boldsymbol{s}) - \sum_{i=1}^{N} \lambda_i f_i(s_i)\Big(J_{\boldsymbol{\lambda}}^*(\boldsymbol{x},t) - J_{\boldsymbol{\lambda}}^*(\boldsymbol{x}+\boldsymbol{e}_i, t)\Big) \right],
$$
(45)

with boundary condition $J_{\boldsymbol{\lambda}}^*(\boldsymbol{x},0) = -\sum_{i=1}^{N} c_m(x_i)$ for all $\boldsymbol{x} \in \mathscr{X}^N$.

To see the reasoning behind (44), consider $N = 2$, $(\boldsymbol{x},t)$ with $x_1, x_2 < \xi$ and $t > 0$, and $\boldsymbol{s} = (s_1, s_2)$. There are four events that can occur:

1. Both system 1 and 2 deteriorate and we move to $J_{\boldsymbol{\lambda}}^*((x_1+1, x_2+1), t-\delta t)$. This happens with probability $\lambda_1 f_1(s_1)\delta t \cdot \lambda_2 f_2(s_2)\delta t = o(\delta t)$;

2. Both system 1 and 2 do not deteriorate and we move to $J_{\boldsymbol{\lambda}}^*((x_1, x_2), t-\delta t)$. This happens with probability $(1 - \lambda_1 f_1(s_1)\delta t) \cdot (1 - \lambda_2 f_2(s_2)\delta t) = 1 - \lambda_1 f_1(s_1)\delta t - \lambda_2 f_2(s_2)\delta t - \lambda_1 f_1(s_1)\delta t \cdot \lambda_2 f_2(s_2)\delta t = 1 - \big((\lambda_1 f_1(s_1) + \lambda_2 f_2(s_2))\big)\delta t + o(\delta t)$;

3. System 1 deteriorates but system 2 does not and we move to $J_{\boldsymbol{\lambda}}^*((x_1+1, x_2), t-\delta t)$. This happens with probability $\lambda_1 f_1(s_1)\delta t \cdot (1 - \lambda_2 f_2(s_2)\delta t) = \lambda_1 f_1(s_1)\delta t - \lambda_1 f_1(s_1)\delta t \cdot \lambda_2 f_2(s_2)\delta t = \lambda_1 f_1(s_1)\delta t + o(\delta t)$;

4. System 2 deteriorates but system 1 does not and we move to $J_{\boldsymbol{\lambda}}^*((x_1, x_2+1), t-\delta t)$. This happens with probability $\lambda_2 f_2(s_2)\delta t \cdot (1 - \lambda_1 f_1(s_1)\delta t) = \lambda_2 f_2(s_2)\delta t - \lambda_2 f_2(s_2)\delta t \cdot \lambda_1 f_1(s_1)\delta t = \lambda_2 f_2(s_2)\delta t + o(\delta t)$.

Event 1 is $o(\delta t)$, event 2 is captured in the "no deterioration" term in in (45), and event 3 and 4 are captured in the "system i deteriorates" term in (45). For general N, the derivation of (44) is along similar lines. Note that for $N = 1$, Equation (45) reduces to the single system HJB equation that we formulated in Lemma 1.

Solving (45) can be done numerically via (a multivariate version of) the solution approach outlined in Appendix C. However, even for moderate values of N, ξ, and $\bar{s}$ (the maximum production rate), the size of the state and action space becomes very large; consequently, obtaining the solution of the HJB recursion (45) is computationally demanding and can only be done for small instances. Hence, in the following section, we will confine our numerical experiments to settings that involve only two production systems.

B.2. Numerical Experiments with Two Production Systems

We first consider two symmetric production systems. Both production systems have the same values as the example in Figure 1. Thus, both systems have $\xi = 10$, $c_m(\xi) = 5$, and $c_m(\cdot) = 1$ for all non-failed states. Further, T equals 15, and the rate space for both systems equals $\mathscr{S} = [0,1]$. For the deterioration function, we use $f_1(s) = f_2(s) = s^2$, and we set the base rates λ_1 and λ_2 to 1. We use $D = 1.5$, $p = 1$, and $b = 0$ for the parameters of the revenue function (43). Thus the cumulative production output of both systems must meet a demand rate of 1.5, and any shortfall of that required production rate is penalized at rate 1. Figure 11 presents joint sample paths of the optimal production policy and the corresponding controlled deterioration for system 1 (top) and system 2 (bottom).
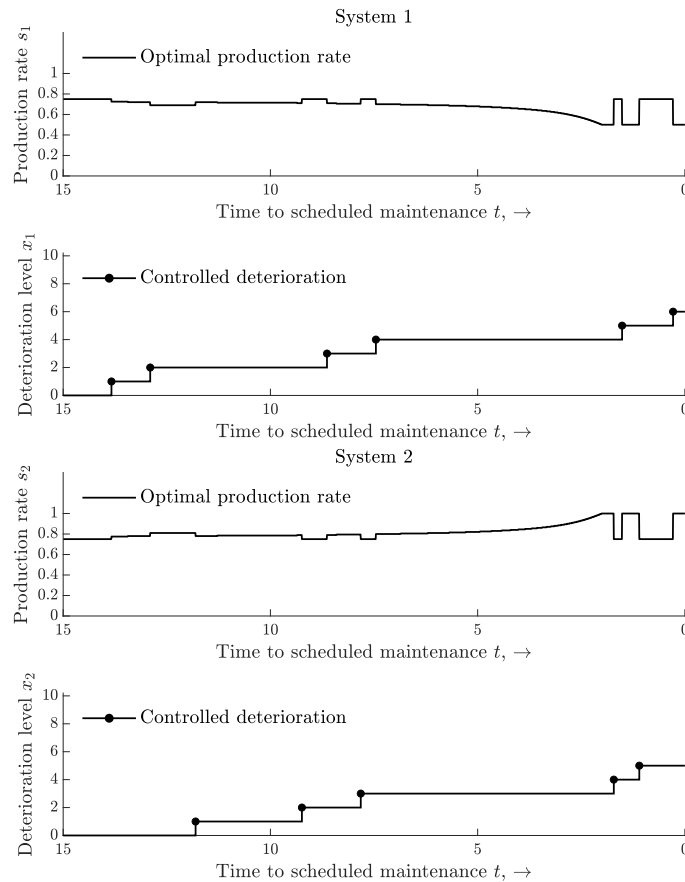
Figure 11   Illustration of the optimal condition-based production policy for two symmetric production systems.

Based on Figure 11, we can state three important observations. First, when both production systems have equal deterioration levels, they share the demand rate requirement equally (i.e. $s_1 = s_2$ when $x_1 = x_2$). Second, if one system's condition worsens, its production rate decreases instantly,

while the other system instantly compensates by an equal increase, thereby still meeting the demand rate requirement. Third, as the time to the planned maintenance moment decreases, the distribution of the required demand rate across both production system changes when the deterioration levels differ. For example, we clearly see that the optimal production rate of system 1 for fixed deterioration level $x_1 = 4$ decreases as t decreases, while the optimal production rate of system 2, which has deterioration level $x_2 = 3$ at that time, increases with the same amount. This latter observation also breaks the monontonic behavior of the optimal production policy in t for a single system (cf. Theorem 1). These three observations together indicate that the optimal condition-based production policy for symmetric production systems leads to synchronized controlled deterioration.

Next, we consider two production systems that are asymmetric in their base rates. We use the same values as the symmetric setting, except for the base rates which are set at $\lambda_1 = 1.5$ and $\lambda_2 = 0.5$, so that production system 1 deteriorates faster under normal operating conditions than production system 2. Figure 12 presents joint sample paths of the optimal production policy and the corresponding controlled deterioration for system 1 (top) and system 2 (bottom).

Figure 12 suggests that the optimal condition-based production policy for asymmetric production systems also results in synchronized controlled deterioration. However, due to differing base rates, the distribution of the demand rate requirement across the two systems varies and is inversely proportional to these base rates. At the start of the planning horizon, we still see that if one system's condition worsens, its production rate decreases instantly, while the other system instantly compensates by an equal increase, thereby still meeting the demand rate requirement. However, as the time to planned maintenance shortens and both systems approach their failure levels, this compensatory effect becomes less pronounced. Notably, close to the end of the planning horizon, the production systems do not even meet the demand rate requirement any more, indicating that the expected maintenance costs have exceeded those associated with the production shortfall.

Appendix C: Solving Continuous Time Dynamic Programs

To compute the optimal policy we discretize the planning horizon $[T, 0]$ into $N \triangleq T/\Delta t$ time intervals of length Δt and use a finite difference equation to approximate the optimality equation in (2). Specifically, we replace the partial derivative in Equation (2) by the finite difference $\big(J_\lambda^*(x, t + \Delta t) - J_\lambda^*(x, t)\big)/\Delta t$, after which we obtain the finite difference equation:

$$
J_\lambda^*(x, t + \Delta t) =
\begin{cases}
J_\lambda^*(x, t) + \Delta t \cdot \max_{s \in \mathscr{S}} \Big[r(s) - \lambda f(s)\big(J_\lambda^*(x, t) - J_\lambda^*(x + 1, t)\big) \Big], & \text{if } x \in \{0, 1, \ldots, \xi - 1\}, \\
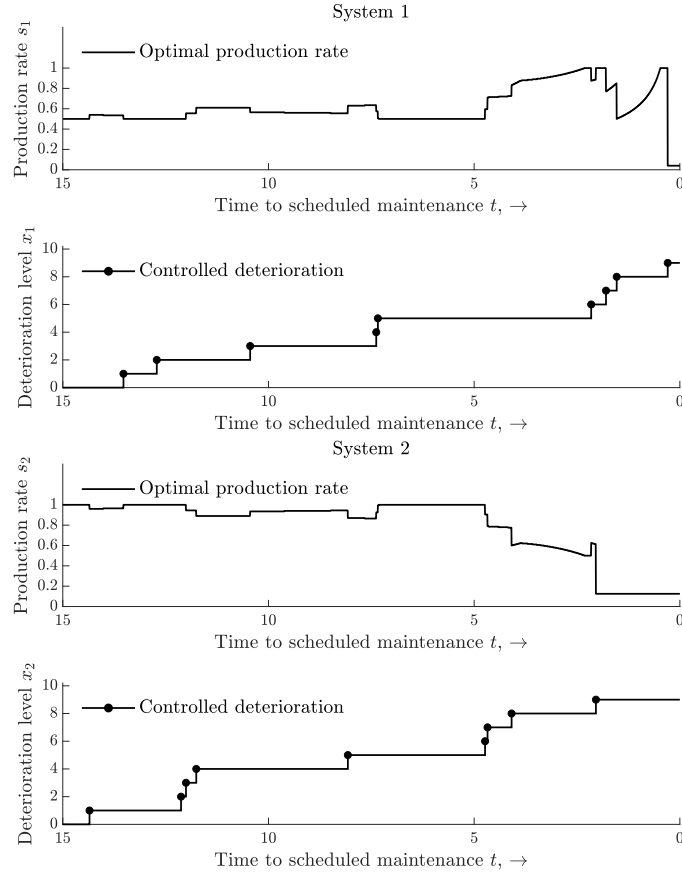J_\lambda^*(x, t), & \text{if } x = \xi,
\end{cases}
$$

Figure 12 Illustration of the optimal condition-based production policy for two asymmetric production systems. System 1 (top) has base rate $\lambda_1 = 1.5$ and system 2 (bottom) has base rate $\lambda_2 = 0.5$.

which can be solved by a backward recursion on the discrete time set $\left\{ n\Delta t \mid n \in \{0, 1, \ldots, N\} \right\}$ starting from the boundary condition $J_\lambda^*(x, 0) = -c_m(x)$ for all $x \in \mathscr{X}$. In all our numerical experiments and illustrations we use a sufficiently small Δt, such that the approximation is sufficiently accurate.

Appendix D: Detailed Results Numerical Experiments

This section presents detailed results from our numerical experiments. We first discuss the results from the full factorial experiments conducted on systems with known, fixed base rates. Subsequently, we discuss the results of the full factorial experiment involving systems exhibiting unit-to-unit variability in the base rates.

D.1. Static versus Condition-Based Production

To assess the value of condition-based production over static production policies, we perform an extensive numerical study. This numerical study has the following parameter settings. We use base rates $\lambda \in \{0.5, 0.75, 1, 1.25, 1.5\}$, failure thresholds $\xi \in \{10, 12, 14, 18, 20\}$, and planning horizons $T \in \{10, 20\}$. We use the functions $f = s^\gamma$ and $r = s^\nu$, for the degradation and revenue function,

respectively, and use the interval $\mathscr{S} = [0,2]$ for the production rates to choose from. By using the values $\{0.5, 0.75, 1, 1.33, 2\}$ for γ and ν, we model concave, linear, and convex functions, respectively. We vary the preventive maintenance cost c_p between $\{1,2,3\}$, and choose $c_m(\xi) = 10$ as corrective maintenance cost for the failed state. The complete set of parameter values is summarized in Table 4. We perform a full factorial test bed, which results in 3750 instances.

Table 4   Input parameter values for numerical study

	Input parameter	No. of choices	Values
1	Base rate, λ	5	0.5, 0.75, 1, 1.25, 1.5
2	Preventive maintenance cost, c_p	3	1, 2, 3
3	Failure threshold, ξ	5	10, 12, 14, 18, 20
4	Length of planning horizon, T	2	10, 20
5	Parameter of deterioration function, γ	5	0.5, 0.75, 1, 1.33, 2
6	Parameter of revenue function, ν	5	0.5, 0.75, 1, 1.33, 2

The results of the numerical study are summarized in Table 5. In this table, we present the average and standard deviation of $\mathscr{R}$, and the $P_{\pi_{CS}}$, where we first distinguish between subsets of instances with the same value for a specific input parameter of Table 4, and then present the results for all instances.

Table 5   Relative profit increase (in %) of condition-based production compared to static production.

Input parameter	Value	Mean $\mathscr{R}$	SD $\mathscr{R}$	$P_{\pi_{CS}}$
Base rate, λ	0.5	15.18	11.46	33.74
	0.75	26.82	18.26	24.89
	1	41.58	34.97	19.39
	1.25	72.31	133.78	15.73
	1.5	96.23	425.59	13.16
Preventive maintenance cost, c_p	1	55.94	262.30	22.38
	2	46.11	188.32	21.38
	3	49.23	136.83	20.39
Failure threshold, ξ	10	92.22	427.09	14.65
	12	66.07	129.02	17.70
	14	42.56	44.98	20.59
	18	27.65	21.93	25.81
	20	23.63	17.87	28.16
Length of planning horizon, T	10	47.94	151.18	19.47
	20	52.91	243.17	23.29
Parameter of deterioration function, γ	0.5	67.08	376.91	31.59
	0.75	61.19	173.84	25.54
	1	48.59	168.19	20.75
	1.33	41.02	45.08	16.42
	2	34.25	42.68	12.60
Parameter of revenue function, ν	0.5	28.95	48.01	12.14
	0.75	32.31	43.44	13.35
	1	53.79	169.41	15.83
	1.33	67.49	175.60	21.49
	2	69.58	374.14	44.09
Total		50.42	202.46	21.38

D.2. Sequential versus Integrated Maintenance Scheduling

To assess the value of integrating maintenance and production policies, we consider a full factorial test bed using the same parameters of the previous full factorial experiment as displayed in Table 4 (excluding the parameter T for obvious reasons), which results in 1875 instances. We summarize the results of this numerical study in Table 6. Again, we present the average and standard deviation of $\hat{\mathscr{R}}$, and the $\hat{P}_{\pi_I}$, where we first distinguish between subsets of instances with the same value for a specific input parameter of Table 4, and then present the results across all instances.

Table 6   Relative profit rate increase (in %) of integrating maintenance and production decisions compared to treating them sequentially.

Input parameter	Value	Mean $\hat{\mathscr{R}}$	SD $\hat{\mathscr{R}}$	$\hat{P}_{\pi_I}$
	0.5	26.18	30.00	2.04
	0.75	23.22	28.99	1.92
Base rate, λ	1	20.80	28.09	1.82
	1.25	19.04	27.29	1.73
	1.5	17.70	26.59	1.64
	1	18.74	24.84	2.04
Preventive maintenance cost, c_p	2	21.44	28.38	1.82
	3	23.99	31.25	1.63
	10	15.95	23.72	1.63
	12	18.39	25.87	1.75
Failure threshold, ξ	14	20.80	27.75	1.83
	18	25.05	30.62	1.95
	20	26.76	31.70	1.99
	0.5	2.73	1.98	1.98
	0.75	8.46	4.56	1.92
Parameter of deterioration function, γ	1	17.40	8.99	1.87
	1.33	30.52	18.88	1.78
	2	47.84	47.40	1.60
	0.5	5.70	6.25	1.00
	0.75	11.25	8.79	1.22
Parameter of revenue function, ν	1	18.23	14.24	1.50
	1.33	27.73	23.56	1.99
	2	44.04	47.30	3.43
Total		21.39	28.35	1.83

D.3. Performance of the Certainty-Equivalent Policy

To assess the performance of the certainty-equivalent policy, we use a full-factorial design in which we vary five different input parameters (see Table 7 for a summary). We use failure thresholds $\xi \in \{10, 15, 20\}$ and planning horizons $T \in \{10, 15, 20\}$. Again, we use a single corrective maintenance cost equal to 10 and a single preventive maintenance cost for all non-failed states, denoted with c_p, which we vary between two values 1 and 5. We use the functions $f = s^{\gamma}$ and $r = s^{\nu}$, for the degradation and revenue function, respectively, and use the interval $\mathscr{S} = [0, 2]$ to choose the production rates from. By using the values $\{0.5, 1, 2\}$ for γ and ν, we model concave, linear, and convex functions, respectively. The certainty-equivalent policy is characterized by the parameter N^{opt}, which we vary between 5 values. In a simulation instance (we shortly describe the simulation procedure), the base rates are sampled from a Gamma distribution in which we fix the mean to 1,

and vary its coefficient of variation, denoted with cv_Λ. This coefficient of variation is thus a measure for the population heterogeneity of the systems. These choices lead to a total of 4050 instances.

Table 7   Input parameter values for simulation study

	Input parameter	No. of choices	Values
1	Coefficient of variation of prior Gamma distribution, cv_Λ	3	$\sqrt{\frac{1}{2}}$, 1, $\sqrt{2}$
2	Preventive maintenance cost, c_p	2	1, 3
3	Length of planning horizon, T	3	10, 15, 20
4	Failure threshold, ξ	5	10, 12, 15, 18, 20
5	Parameter of deterioration function, γ	3	0.5, 1, 2
6	Parameter of revenue function, ν	3	0.5, 1, 2
7	Number of re-optimizations, N^{opt}	5	1, 2, 4, 6, 8

The results of the simulation study are summarized in Table 8. In this table, we present the average and maximum $\bar{\mathscr{R}}$, and the $\bar{P}_{\pi_O}$, where we first distinguish between subsets of instances with the same value for a specific input parameter of Table 7, and then present the results across all instances.

Table 8   Results of simulation study: Relative regret (profit loss in %) compared with the Oracle performance.

Input parameter	Value	Mean $\bar{\mathscr{R}}$	Max $\bar{\mathscr{R}}$	$\bar{P}_{\pi_O}$
Coefficient of variation of prior Gamma distribution, cv_Λ	$\sqrt{\frac{1}{2}}$	0.59	7.29	10.32
	1	0.72	9.75	10.33
	$\sqrt{2}$	0.84	11.63	10.33
Preventive maintenance cost, c_p	1	0.65	10.06	11.33
	5	0.78	11.63	9.33
Length of planning horizon, T	10	0.38	3.10	7.03
	15	0.70	7.16	10.51
	20	1.08	11.63	13.46
Failure threshold, ξ	10	1.37	11.63	9.03
	12	0.94	7.98	9.75
	15	0.64	4.98	10.32
	18	0.34	2.33	11.14
	20	0.28	2.09	11.43
Parameter of deterioration function, γ	0.5	0.01	0.07	9.93
	1	0.78	11.63	10.24
	2	1.39	11.63	10.83
Parameter of revenue function, ν	0.5	1.39	11.63	10.83
	1	0.77	11.63	10.24
	2	0.01	0.05	9.93
Number of re-optimizations, N^{opt}	1	0.91	11.63	10.32
	2	0.79	10.21	10.33
	4	0.68	8.64	10.33
	6	0.62	7.90	10.33
	8	0.59	7.55	10.33
Total		0.72	11.63	10.33

Table 8 indicates that the certainty-equivalent policy performs excellently with average profit losses of only 0.72% on average and a maximum profit loss of only 11.63% over all 4050 instances. The instance with $cv_\Lambda = \sqrt{2}$, $c_p = 5$, $T = 20$, $\xi = 10$, $\gamma = 1$ or 2, $\nu = 0.5$ or $\nu = 1$ and $N^{\mathrm{opt}} = 1$ leads to this maximum profit loss of 11.63%. As such, the value 11.63 occurs as the max $\bar{\mathscr{R}}$ for all subsets of instances containing this poorly performing instance in Table 8. Indeed, we observe that if you

deviate from these values and hence exclude the instance with the highest max $\bar{\bar{\mathscr{R}}}$, the max $\bar{\bar{\mathscr{R}}}$ of the resulting subset of instances is smaller. This is the case, when $cv_\lambda < \sqrt{2}$, $c_p = 5$, $T < 20$, $\xi > 10$, $\gamma = 0.5$, $\nu = 2$, or $N^{\mathrm{opt}} > 1$.

Some insightful observations can be drawn from Table 8. We find that the profit loss is inversely related to the number of re-optimizations and the failure threshold. The decrease in the number of re-optimizations is intuitive; the certainty equivalent policy adapts more frequently to the accrued knowledge of the current system's base rate yielding a lower profit loss with respect to the optimal policy for that base rate. When the failure threshold increases there is less need for controlled deterioration as the system can, on average, be set to its maximum production rate more often. This holds for both the Oracle policy and the certainty-equivalent policy resulting in comparable performance. Finally, we see that the form of both the degradation and revenue function has a huge impact on the relative profit loss; a concave deterioration function and/or a convex revenue function lead to negligible profit losses. In light of Corollary 1, these results suggest that the certainty-equivalent policy performs comparable to the Oracle policy in the bang-bang regime, regardless of the degree of unit-to-unit variability. In §6.4, we numerically confirm this intuition.