European Journal of Operational Research 227 (2013) 88–100

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Production, Manufacturing and Logistics

The impact of process deterioration on production and maintenance policies

Burak Kazaz [a,1], Thomas W. Sloan [b,*]

[a] Whitman School of Management, Syracuse University, 721 University Ave., Syracuse, NY 13244, United States
[b] Manning School of Business, University of Massachusetts Lowell, 1 University Ave., Lowell, MA 01854, United States

ARTICLE INFO

Article history:
Received 21 September 2010
Accepted 27 November 2012
Available online 13 December 2012

Keywords:
Manufacturing
Maintenance
Scheduling

ABSTRACT

This paper examines a single-stage production system that deteriorates with production actions, and improves with maintenance. The condition of the process can be in any of several discrete states, and transitions from state to state follow a semi-Markov process. The firm can produce multiple products, which differ by profit earned, expected processing time, and impact on equipment deterioration. The firm can also perform different maintenance actions, which differ by cost incurred, expected down time, and impact on the process condition. The firm needs to determine the optimal production and maintenance choices in each state in a way that maximizes the long-run expected average reward per unit time.

The paper makes four contributions: (1) It introduces three critical ratios for the firm's choices. The first enables the firm to decide whether to manufacture or perform maintenance, the second reveals the best product to manufacture, and the third determines the best maintenance action. The economic interpretations of these critical ratios provide managerial insights. (2) The paper shows how the critical ratios can be combined in order to determine the optimal policy, simultaneously accounting for the trade-offs involving production profits, maintenance costs, and the impact on the process condition. We show how these results generalize to problem settings with an arbitrary number of machine states. (3) The paper demonstrates the impact of market demand conditions on the optimal policy. And (4) it develops a set of sufficient conditions that lead to monotone optimal policies. These conditions generalize those reported in earlier studies.

1. Introduction

In many manufacturing environments, the condition of the process or equipment has a significant impact on the quantity and quality of units produced. Consider the case of semiconductor manufacturing in which a chip maker must decide how to allocate production resources among leading-edge and lagging-edge technology products. High-technology products earn a greater profit than low-technology products, but they are also more complex and thus take more time to produce. This increase in production time causes greater deterioration of the manufacturing process, which, in turn, increases the likelihood of quality problems. The chip maker also has the option of performing maintenance, which returns the process to an improved state. Here, too, there is more than one option. At one end of the spectrum, major improvement in the process condition can be achieved by performing a lengthy and costly maintenance procedure; thus, a major maintenance action has a greater likelihood of improving the process condition. At the other end of the spectrum, a minor maintenance procedure can be performed which will cost less and take less time, but has a smaller probability of returning the process to an improved state. Thus, operating this type of system over time requires a manager to answer a series of interconnected questions:

1. Whether to manufacture a product (which may result in the deterioration of the process) or maintain the equipment for a possible improvement.
2. If the decision is to manufacture a product, then which product to manufacture as the choice influences the deterioration of the process differently.
3. If the decision is to maintain the equipment, then which maintenance action to implement as the choice influences the improvement of the process differently.

This paper presents a semi-Markov decision process model to explore the trade-offs involved in answering these three questions. The objective of the model is to determine a course of action that will maximize the long-run expected average reward.

While there has been much research on production systems with deteriorating process condition, our inclusion of multiple products and multiple maintenance actions, as well as our approach, sets this work apart from the majority of previous research

* Corresponding author. Tel.: +1 978 934 2857; fax: +1 978 934 4034.
E-mail addresses: bkazaz@syr.edu (B. Kazaz), Thomas_Sloan@uml.edu (T.W. Sloan).
[1] Tel.: +1 315 443 7381; fax: +1 315 442 1461.

0377-2217/$ - see front matter
http://dx.doi.org/10.1016/j.ejor.2012.11.052

in this area. After developing initial insights about the structural properties of the problem in a four-state setting, we show how these lessons can be extended to a problem with any number of states. We integrate market demand conditions by enforcing minimum and maximum production requirements for each product, and we explore how these constraints impact the resulting optimal policy.

The paper makes four contributions. First, it develops three types of *critical ratios* which allow the comparison of any two actions in a given state. Following the three questions above, one critical ratio determines whether the firm should produce a product or perform maintenance, another determines which product is optimal in states where production is preferred, and a third critical ratio identifies which maintenance action is optimal in states where maintenance is preferred. The first critical ratio can be interpreted as the *reservation price*, i.e., the minimum amount of money the decision maker should earn in order to justify production over maintenance. Similarly, the second critical ratio represents the minimum amount of profit the firm needs to earn to switch from a low-end product to a high-end product. The third critical ratio establishes an upper bound on the maximum amount of money that the decision maker should be willing to spend on major maintenance. Second, the paper demonstrates how the critical ratios can be combined to determine the optimal action in a particular machine state given all of the possible alternatives. Their combination enables the decision maker to *simultaneously* account for the trade-offs involving profit benefits versus deterioration probability and cost versus improvement probability. Third, the paper shows the influence of minimum and maximum throughput requirements on the choice of the optimal policy. It proves that the frequency and timing of maintenance play a strategic role in increasing the throughput of a high demand product. Fourth, the paper develops a set of conditions which are sufficient to ensure that a monotone policy is optimal. In monotone policies, the firm manufactures the high-end product in better states, and switches to the low-technology product as the process deteriorates. Minor maintenance is employed as the process continues to deteriorate, eventually employing major maintenance at significant deterioration levels. The conditions that lead to monotone policies are much more general than those reported in previous research. We demonstrate the utility of the new conditions by presenting examples that do not meet the previously reported conditions but that still have monotone optimal policies.

The paper proceeds as follows. The next section presents an overview of the relevant literature. The basic model is developed in Section 3. Section 4 examines the impact of adding production requirements. Section 5 presents several generalizations of the model and discusses how our results go beyond those previously reported. Conclusions and managerial insights are in Section 6. All proofs and technical derivations are provided in the Appendix.

2. Literature review

Many researchers have studied problems at the intersection of production and maintenance scheduling, i.e., where the state of the equipment affects the production process in some way. Production systems with variable yield have received and continue to receive much attention, as discussed in the extensive reviews by Yano and Lee (1995) and Hadidi et al. (2012). Much of the work in this area, starting with Rosenblatt and Lee (1986) and Porteus (1986), has been a variation of the economic manufacturing quantity (EMQ) model. The central questions is: How much of a product should be produced given that some fraction of it may be defective? The process begins in an "in-control" state but may shift to an "out-of-control" state, which results in defective products. Groenevelt et al. (1992a,b) and El-Ferik (2008) show that the optimal batch sizes are bigger when the possibility of equipment failure is incorporated. These models account for the risk of unknowingly producing defective items, and the equipment state affects only the quantity of production, but not the quality.

These early works have been extended in many ways. Hariga and Ben-Daya (1998) relax some of the assumptions about the equipment's shift to the out-of-control state and develop structural properties for this more general case. Lee and Rosenblatt (1989) and Lee and Park (1991) investigate different cost structures that depend on when defective items are detected. Lee and Rosenblatt (1987), Porteus (1990), Makis (1998), and Kim et al. (2001) incorporate inspections into the decision model, allowing early detection of the out-of-control state. Boone et al. (2000) extend the model to include machine failures, and Makis and Fung (1998) include both inspections and machine failures. The model proposed by Ben-Daya (2002) allows for imperfect maintenance, i.e., preventive maintenance that may not return the process to the in-control state. Similar models with imperfect maintenance have been proposed by Chakraborty et al. (2008), Liao et al. (2009), and Sana (2010a,b), each making different assumptions about how the process drifts to the out-of-control state, repair times, and yield distributions.

Departing from the EMQ approach, Gilbert and Emmons (1995) develop a model of a job shop in which defective items must be reworked and reduce the production capacity. Inspections reveal if the process is out of control, and a restoration action returns the process to the in-control state. The objective is to determine an inspection and restoration policy that maximizes throughput. Gilbert and Bar (1999) extend these ideas to a small batch production system where they show that a control limit policy is optimal, suggesting that it is ideal to restore the equipment condition when the number of units remaining in a batch exceed a certain threshold.

Sloan (2004) models a system with multiple machine states, where the output follows a binomial distribution that depends on the equipment state. Iravani and Duenyas (2002) construct an integrated production and maintenance model in which the decisions at each epoch are restricted to: produce one unit (rather than in batches), perform maintenance, or do nothing.

While the papers mentioned above consider single-product systems, a significant amount of research has investigated multi-product systems. For example, Lee (2004) examines a traditional job-shop scheduling problem in the context of unreliable equipment. Cassady and Kutanoglu (2005) extend this type of work by simultaneously determining the maintenance and production schedules. Aghezzaf et al. (2007) also aim to combine production and maintenance scheduling, this time in the context of a multi-product, batch production system with failure-prone equipment. Karamatsoukis and Kyriakidis (2010) examine a two-stage production system, where the two stages deteriorate independently and are separated by an inventory buffer. The maintenance policy — including preventive and corrective maintenance — is influenced by the inventory level. Dehayem Nodem et al. (2009) also examine a system with maintenance actions including imperfect repair and replacement. They use a semi-Markov decision framework to determine the optimal production rate and maintenance policy. Nourelfath (2011) studies a multi-period, multi-product production system in which the maintenance policy affects product availability. A constrained, stochastic capacitated lot-sizing approach is used to ensure a given service level. In all of these papers as well, however, the state of the process is limited to either "up" or "down." In the "down" state, no production is possible; in the "up" state, all output is of perfect quality.

Sloan and Shanthikumar (2000, 2002) study multi-product systems with deteriorating process condition in which the process state can be influenced by the decision maker and where the state affects the yield of each product differently. However, both studies assume that all products have the same processing times and the machine state transitions are independent of the product manufactured. Batun and Maillart (2012) use a similar framework to examine different dispatching policies. Kazaz and Sloan (2008) consider a single-stage system in which processing times and machine state transition probabilities both vary by product type. Conditions are developed that define the exact optimality point for each product and state; however, no demand requirements are considered. In addition, in all of these papers only one maintenance action is allowed, and this action returns the process to the best state with probability one.

In most situations, there are maintenance actions short of total replacement that can be taken to reduce or alter the rate of process deterioration. Wang (2002) provides an extensive review of the maintenance literature. The works that relate most closely to the current problem include single-machine systems with Markov deterioration and multiple maintenance actions. Such models have been formulated in the context of completely observable state information (Hopp and Wu, 1990), partially observable state information (Hopp and Wu, 1988), and imperfect maintenance (Su et al., 2000). None of these models, however, explicitly accounts for the impact of equipment condition on the production process.

In sum, there has been relatively little work on systems that have the following characteristics: multiple products are produced, the quality of output depends on the equipment or process state, the process state can be influenced by the decision maker, and multiple "maintenance" actions are allowed. One model that addresses all of these issues — and therefore most closely relates to ours — is that of Sloan (2008), which studies a multi-product manufacturing system in which multiple maintenance actions are available. The processing times and associated machine state transition probabilities both depend on the type of production and maintenance actions being employed. Sufficient conditions are developed that ensure a monotone policy with respect to both production and maintenance actions. Our paper also provides sufficient conditions; however, the ones presented here are significantly more general than those presented in Sloan (2008). We demonstrate the utility of the new conditions by presenting example problems with monotone optimal policies that do not meet the conditions of Sloan (2008) but do meet the new set of conditions.

3. The model

This section presents a model to determine a firm's production and maintenance decisions in a single-stage manufacturing process. The equipment used in the process is described by a discrete number of states denoted by $i = 1, \ldots, N$. As the equipment condition deteriorates, state i moves from 1 (best state) to N (worst state). The equipment condition deteriorates as production takes place and improves with maintenance. The firm is capable of producing multiple products, where each product influences the deterioration process differently. We denote $P1$ as a standard, low-end technology product and $P2$ as a new, high-end technology product. Similarly, the firm can take various maintenance actions that result in varying improvements in the state of the equipment. We denote $M1$ as a minor maintenance action and $M2$ as a major maintenance action. We define the set of production decisions as $\mathbf{P} = \{P1, P2\}$ and the set of maintenance decisions as $\mathbf{M} = \{M1, M2\}$. The firm's objective is to determine a course of action that maximizes the long-run expected average reward. As a result, the manager is faced with the following three decisions at each decision epoch: (1) whether to manufacture a product or perform maintenance, (2) if production is picked, then which product to produce, and (3) if maintenance is picked, then which type of maintenance to perform.

Each of the above three decisions has trade-offs for the manufacturer. In the case of the first decision, the firm has to choose between manufacturing and maintenance actions. When manufacturing is the choice, the firm earns a profit via its production but risks the deterioration of the equipment further. However, when maintenance is the choice, the firm incurs a cost for maintaining the system (rather than earning profit) but increases the likelihood of improving the equipment condition. In addition, more time spent maintaining the equipment means less time producing, so while the improved equipment condition will increase the yield, the net throughput may actually decrease.

We define a_i as the action taken in state $i = 1, \ldots, N$ which consists of manufacturing choices such as $P1$ and $P2$ and maintenance choices such as $M1$ and $M2$; thus $a_i \in \{P1, P2, M1, M2\}$ for all $i = 2, \ldots, N-1$. In order to reflect the operating environment of a manufacturer, we require that the firm manufactures in the best state, i.e., $a_1 \in \{P1, P2\}$, and that it performs maintenance in the worst state, i.e., $a_N \in \{M1, M2\}$. State transition probabilities depend on the choices of manufacturing and maintenance actions. We define $p_{ij}^{a_i}$ as the probability that the machine moves from state $i = 1, \ldots, N$ to state $j = 1, \ldots, N$ when action a_i is taken in state i. When a manufacturing action is taken in state i, the equipment either stays in its current state or deteriorates to a worse state, but cannot improve to a better state. In other words, $p_{ij}^{a_i \in \mathbf{P}} > 0$ for all $i \leqslant j$ where $i = 1, \ldots, N-1$ and $j = i, \ldots, N$, and $p_{ij}^{a_i \in \mathbf{P}} = 0$ for all $i > j$ where $i = 1, \ldots, N-1$ and $j = 1, \ldots, i-1$. On the other hand, when a maintenance action is taken in a state i, the equipment either stays in its current state or improves to a better state, but cannot deteriorate to a worse state. Thus, $p_{ij}^{a_i \in \mathbf{M}} > 0$ for all $i \geqslant j$ where $i = 2, \ldots, N$ and $j = 1, \ldots, i$, and $p_{ij}^{a_i \in \mathbf{M}} = 0$ for all $i < j$ where $i = 1, \ldots, N-1$ and $j = i+1, \ldots, N$. For any state i where $1 < i < N$, the machine state transition probabilities can be summarized as follows:

$$p_{ij}^{a_i} \begin{cases} = 0 \text{ when } j < i; \quad > 0 \text{ when } j > i \quad \text{for } a_i = P1, P2 \quad \text{where } p_{ij}^{P1} < p_{ij}^{P2}; \\ > 0 \text{ when } j = i; \quad \text{and } p_{ii}^{P1} < p_{ii}^{P2}; \quad \text{and } p_{ii}^{M1} > p_{ii}^{M2}; \\ = 0 \text{ when } j > i; \quad > 0 \text{ when } j < i \quad \text{for } a_i = M1, M2 \quad \text{where } p_{ij}^{M1} < p_{ij}^{M2}. \end{cases} \tag{1}$$

Regarding the choice of product to be manufactured, the firm has to consider another trade-off as well. In this case, the firm needs to decide whether to earn a regular profit with a lower risk of deterioration versus a higher profit that comes with an increased likelihood of deterioration. The profit earned from each product is denoted as ρ_{a_i} where $a_i \in \mathbf{P}$. As consumers are willing to pay more for a new technology item and less for a standard product, we assume that the standard product earns a smaller profit than the new product; i.e. $\rho_{P1} < \rho_{P2}$. The yield from manufacturing activities also vary by product and by state. We define y_{i,a_i} as the amount of yield for product $a_i \in \mathbf{P}$ when manufactured in state i, and assume that y_{i,a_i} is decreasing in state i as the firm obtains a lower number of non-defective products with deteriorating process conditions. We define the total profit generated in state i by production action $a_i \in \mathbf{P}$ as $r_{i,a_i} = \rho_{a_i} y_{i,a_i}$. In a typical operating environment for a semiconductor manufacturer, the high-end product generates a larger total profit in each state, i.e., $r_{iP1} < r_{i,P2}$ in each state $i = 1, \ldots, N-1$. However, the processing times also vary by product and by state, and can make the manufacturing of the high-end product less desirable. We define the expected processing time for these two production choices in a state i as $\tau_{i,P1}$ and $\tau_{i,P2}$, respectively. In semiconductor manufacturing, new products typically require a higher circuit density and have a longer expected processing time than the older products. Reflecting this fact, we assume $\tau_{i,P2} > \tau_{i,P1}$ in each state $i = 1, \ldots, N-1$. The consequence of a longer expected processing time is that the equipment is more likely to deteriorate when product $P2$ is manufactured. Thus, it is appropriate to define the state transition probabilities as $p_{ii}^{P1} > p_{ii}^{P2}$ corresponding to the fact that the equipment would stay in its current state with a higher probability when product $P1$ is produced than when product $P2$ is produced. Alternatively for state i, the firm has $\sum_{j=i+1}^{N} p_{ij}^{P1} = (1 - p_{ii}^{P1}) < \sum_{j=i+1}^{N} p_{ij}^{P2} = (1 - p_{ii}^{P2})$, and the sum of deterioration probabilities is lower when product $P1$ is produced than when product $P2$ is manufactured. It should be emphasized here that the decreasing behavior of $r_{i,a_i \in \mathbf{P}}$ and the increasing behavior of $\tau_{i,a_i \in \mathbf{P}}$ are not necessary in developing our results in Section 3. However, they represent the operating environment for semiconductor manufacturers, and more importantly, are useful in explaining the structural results regarding monotone optimal policies in Section 5.

The third question considers the trade-off in alternative maintenance actions. The standard maintenance action $M1$ has a cost of $c_{i,M1} > 0$ and its expected processing time in state i is defined as $\tau_{i,M1}$. In this case, the firm can take a more involved maintenance action described by $M2$. The cost of maintenance action $M2$ is higher than that of $M1$: $c_{i,M2} > c_{i,M1}$ for all states $i = 2, \ldots, N$. However, the likelihood of improving the process condition through $M2$ is also higher. Thus, the firm has $p_{ii}^{M2} < p_{ii}^{M1}$ and the sum of improvement probabilities are $\sum_{j=1}^{i-1} p_{ij}^{M2} = (1 - p_{ii}^{M2}) > \sum_{j=1}^{i-1} p_{ij}^{M1} = (1 - p_{ii}^{M1})$. It is assumed that the major maintenance action $M2$ has a longer expected processing time than the minor maintenance action $M1$ in each state, and therefore $\tau_{i,M2} > \tau_{i,M1}$ in each state $i = 2, \ldots, N$.

Process deterioration can also influence the cost of maintenance. For example, as the equipment deteriorates more, the firm might have to spend more effort and money on maintenance. Therefore, $c_{i,a_i \in \mathbf{M}}$ and $\tau_{i,a_i \in \mathbf{M}}$ can be considered as increasing in i. Once again, the increasing behavior of $c_{i,a_i \in \mathbf{M}}$ and $\tau_{i,a_i \in \mathbf{M}}$ are not necessary for the results developed in Section 3. However, they prove to be useful in explaining the conditions that lead to monotone optimal policies in Section 5.

Note that the state transition probabilities for maintenance actions are defined in a more general way than in most previous research. For example, most research assumes that maintenance returns the machine to the best state with probability one; we make no such assumption. In addition, the transition probability from state i to j where $i > j$ when maintenance action $a_i \in \mathbf{M}$ is taken in state i need not be equal to the transition probability from state k to j where $k > j$ when the same maintenance action is taken in state $k \neq i$. Nor does the paper make an assumption such as $p_{ij}^{a_i} = p_{kj}^{a_i}$ where $\{i,k\} > j$ for each maintenance action $a_i \in \mathbf{M}$. Moreover, this paper does not assume that improvement probabilities with a constant number of states are equal. For example, $p_{ij}^{a_i}$ and $p_{i+1,j+1}^{a_i}$ where $i > j$ are not necessarily equal for a maintenance action $a_i \in \mathbf{M}$.

It should be observed that the time between decisions epochs, the state transition probabilities, the profits and the maintenance costs are dependent only on the action taken in the current state. Therefore, the problem can be modeled as a Semi-Markov Decision Process (SMDP). A time-invariant (or, stationary) policy results in a discrete-time Markov chain that represents the machine condition at decision epochs, and is referred to as the *Embedded Markov Chain* (EMC). The state transition probabilities in this problem characterize the evolution of the EMC over time as they can be defined as $p_{ij}^{a_i} = \Pr\{X_{t+1} = j | X_t = i, a_t = a\}$ where X_t denotes the machine state and a_t describes the action taken at decision epoch t. While there are several approaches to solving this type of problem (interested readers can review Puterman, 1994), we utilize a *policy improvement* approach. In this approach, we begin with a reference policy and compare it to other policies that differ in its actions in various states. The policy that maximizes the long-run average expected value is referred to as the optimal policy.

We define $\mathbf{A} = [a_i | i = 1, \ldots, N]$ as a stationary policy that describes the firm's action a_i in state i and $\Pi_i(\mathbf{A})$ corresponds to the steady-state probability that the EMC is in state i. It should be observed that given the definition of the state transition probabilities, the EMC induced by a stationary policy $\mathbf{A}$ has a single closed set of recurrent states (i.e., is unichain). The implication of having a single set of recurrent states is that, regardless of the initial state of the equipment, there exists a unique set of steady-state probabilities. However, the steady-state probability, defined as $\Pi_i(\mathbf{A})$ for state i, depends on the actions taken in all states. Because the profits and costs depend only on the actions taken in the current state, $EV(\mathbf{A}) = \sum_{i=1}^{N} (1_{a_i \in \mathbf{P}} r_{i,a_i} \Pi_i(\mathbf{A}) - 1_{a_i \in \mathbf{M}} c_{i,a_i} \Pi_i(\mathbf{A})) / \sum_{i=1}^{N} (\tau_{i,a_i} \Pi_i(\mathbf{A}))$ is the average reward rate of policy $\mathbf{A}$, where $1_{a_i \in \mathbf{P}}$ is the indicator whether a production action is taken in state i and $1_{a_i \in \mathbf{M}}$ indicates whether a maintenance action is taken. A policy is the optimal policy, described as $\mathbf{A}^*$, when $EV(\mathbf{A}^*) \geqslant EV(\mathbf{A})$ for all stationary policies $\mathbf{A}$. The optimal action in state i is defined as a_i^*, and it can be shown that the optimal policy specifies only one action per state (Puterman, 1994).

The problem variant with four machine states ($i = 1, 2, 3, 4$), two products ($P1$, $P2$), and two maintenance actions ($M1$, $M2$) is sufficient to develop the insight necessary for the structural properties. While Section 3 analyzes the problem with four states, its results are generalized by considering an arbitrary number of states in Section 5. In the four-state variant of the problem, the firm manufactures in the best state, i.e., $a_1 \in \mathbf{P}$, and performs maintenance in the worst state, i.e., $a_4 \in \mathbf{M}$. In the intermediate states ($i = 2, 3$), the firm has to determine an answer to all three questions described earlier: (1) whether to manufacture a product or maintain the equipment, (2) if manufacturing is preferred, then, which product to produce, and (3) if maintenance is preferred, then whether to employ a minor or major maintenance action; thus, $(a_2, a_3) \in \{P1, P2, M1, M2\}$. This results in four groups of policies that feature production and maintenance actions, described with P and M, respectively. These policies are classified as Group 1: [**P**, **P**, **P**, **M**], Group 2: [**P**, **P**, **M**, **M**], Group 3: [**P**, **M**, **P**, **M**], and Group 4: [**P**, **M**, **M**, **M**]. As a result, the firm has a total of 64 policies, as shown in Table 1.

A comparison of the steady-state probabilities in the four groups of policies provides useful observations. We express the steady-state probability in state i as $\Pi_i(\mathbf{A}_n) = \widehat{\Pi}_i(\mathbf{A}_n) / \sum_{i=1}^{N} \widehat{\Pi}_i(\mathbf{A}_n)$, where $\widehat{\Pi}_i(\mathbf{A}_n)$ is the numerator term for the steady-state expression of state i for policy $\mathbf{A}_n$. The $\widehat{\Pi}_i(\mathbf{A}_n)$ values for the 64 policies above can be expressed as

$$\widehat{\Pi}_1(\mathbf{A}_n) = \left\{ \begin{array}{ll} (1 - p_{22}^{a_2})(1 - p_{33}^{a_3}) p_{41}^{a_4} & \text{for } n = 1, \ldots, 16; \\ (1 - p_{22}^{a_2}) p_{31}^{a_3} (p_{41}^{a_4} + p_{42}^{a_4}) + p_{23}^{a_2} p_{31}^{a_3} p_{42}^{a_4} + p_{24}^{a_2} p_{32}^{a_3} p_{41}^{a_4} & \text{for } n = 17, \ldots, 32; \\ (1 - p_{22}^{a_2})(1 - p_{33}^{a_3})(p_{41}^{a_4} + p_{42}^{a_4}) & \text{for } n = 33, \ldots, 48; \\ (1 - p_{22}^{a_2})(1 - p_{33}^{a_3})(1 - p_{44}^{a_4}) & \text{for } n = 49, \ldots, 64. \end{array} \right\} \tag{2}$$

$$\widehat{\Pi}_2(\mathbf{A}_n) = \left\{ \begin{array}{ll} (1 - p_{33}^{a_3}) [p_{12}^{a_1} p_{41}^{a_4} + (1 - p_{11}^{a_1}) p_{42}^{a_4}] & \text{for } n = 1, \ldots, 16 \quad \text{and } n = 33, \ldots, 48; \\ (1 - p_{33}^{a_3}) [p_{12}^{a_1} (1 - p_{44}^{a_4}) + p_{14}^{a_1} p_{42}^{a_4}] + p_{32}^{a_3} [p_{13}^{a_1} (1 - p_{44}^{a_4}) + p_{14}^{a_1} p_{43}^{a_4}] & \text{for } n = 17, \ldots, 32; \ n = 49, \ldots, 64. \end{array} \right\} \tag{3}$$

$$\widehat{\Pi}_3(\mathbf{A}_n) = \begin{Bmatrix} p_{13}^{a_1}(1-p_{22}^{a_2})p_{41}^{a_4} + p_{23}^{a_2}[p_{12}^{a_1}p_{41}^{a_4} + (1-p_{11}^{a_1})p_{42}^{a_4}] + (1-p_{11}^{a_1})(1-p_{22}^{a_2})p_{43}^{a_4} & \text{for } n = 1,\ldots,32; \\ (1-p_{22}^{a_2})[p_{13}^{a_1}(1-p_{44}^{a_4}) + p_{14}^{a_1}p_{43}^{a_4}] & \text{for } n = 33,\ldots,64. \end{Bmatrix} \tag{4}$$

$$\widehat{\Pi}_4(\mathbf{A}_n) = \begin{Bmatrix} (1-p_{11}^{a_1})(1-p_{22}^{a_2})(1-p_{33}^{a_3}) & \text{for } n = 1,\ldots,16; \\ p_{14}^{a_1}(1-p_{22}^{a_2})p_{31}^{a_3} + p_{24}^{a_2}[p_{12}^{a_1}p_{31}^{a_3} + (1-p_{11}^{a_1})p_{32}^{a_3}] & \text{for } n = 17,\ldots,32; \\ (p_{13}^{a_1} + p_{14}^{a_1})(1-p_{22}^{a_2})(1-p_{33}^{a_3}) & \text{for } n = 33,\ldots,48; \\ p_{14}^{a_1}(1-p_{22}^{a_2})(1-p_{33}^{a_3}) & \text{for } n = 49,\ldots,64. \end{Bmatrix} \tag{5}$$

As can be seen from above, the steady-state probability for a state differs from one policy to another, complicating the evaluation of the expected value gained from each policy.

The analyses in Sections 3.1, 3.2, 3.3 investigate the firm's preferred action in a deteriorated intermediate state, specifically state 3. In order to develop insight into the actions in an intermediate state, we restrict our analysis to the case where the actions in states 1 and 2 are $P2$ and $P1$, respectively, and the action in state 4 is limited to the standard maintenance action $M1$. This setting enables us to investigate the impact of all four actions available in state 3, i.e., $a_3 \in \{P1, P2, M1, M2\}$. As a result, the analysis in these sections is restricted to choosing between four policies: $\mathbf{A}_9 = [P2, P1, P1, M1]$, $\mathbf{A}_{11} = [P2, P1, P2, M1]$, $\mathbf{A}_{25} = [P2, P1, M1, M1]$, and $\mathbf{A}_{27} = [P2, P1, M2, M1]$. We begin the discussion with the firm's first decision corresponding to whether to produce or maintain the equipment in an intermediate state.

3.1. The choice between production and maintenance in an intermediate state

As the firm manufactures in states 1 and 2, it has to determine whether it should continue to produce when the process deteriorates to state 3, or alternatively, maintain it in the hope that it returns to better states (1 and 2). We develop a critical ratio of the total profit earned from the manufacturing action (profit per unit times the yield) with respect to the maintenance cost in intermediate states. This critical ratio enables the firm to determine which action, production or maintenance, is a better alternative for the state in question. To see this, we compare the following two policies: $\mathbf{A}_9 = [P2, P1, P1, M1]$ and $\mathbf{A}_{25} = [P2, P1, M1, M1]$. The firm alters its decision only in the third state in these two policies. It is known from (2), (3), and (5) that the steady-state probabilities for states 1, 2 and 4 are different for these two policies despite the fact that they feature the same actions. Similarly, from (4), the steady-state probability for state 3 is also different, and one cannot readily tell whether their relative values increase or decrease. We define $\Delta_{i,j}^{M1,P1}$ as the change in the numerator term of state i when the firm switches from implementing the maintenance action $M1$ to manufacturing product $P1$ in state j. The relationship between the numerator terms are expressed as $\widehat{\Pi}_1(\mathbf{A}_9) = \widehat{\Pi}_1(\mathbf{A}_{25}) - \Delta_{1,3}^{M1,P1}$, where $\Delta_{1,3}^{M1,P1} < \min\{\widehat{\Pi}_1(\mathbf{A}_{25}), \widehat{\Pi}_3(\mathbf{A}_{25})\}$; $\widehat{\Pi}_2(\mathbf{A}_9) = \widehat{\Pi}_2(\mathbf{A}_{25}) - \Delta_{2,3}^{M1,P1}$, where $\Delta_{2,3}^{M1,P1} < \min\{\widehat{\Pi}_2(\mathbf{A}_{25}), \widehat{\Pi}_3(\mathbf{A}_{25})\}$; $\widehat{\Pi}_3(\mathbf{A}_9) = \widehat{\Pi}_3(\mathbf{A}_{25})$; and $\widehat{\Pi}_4(\mathbf{A}_9) = \widehat{\Pi}_4(\mathbf{A}_{25}) + \Delta_{4,3}^{M1,P1}$, where $\Delta_{2,3}^{M1,P1} < \min\{\widehat{\Pi}_3(\mathbf{A}_{25}), \widehat{\Pi}_4(\mathbf{A}_{25})\}$. Using these expressions, the decision maker can develop a critical ratio that determines her preference in state 3.

Proposition 1. *There exists a critical ratio that determines the firm's choice between manufacturing and maintenance in state 3:*

$$\begin{aligned}\gamma_3^{M1,P1} = &-\left(\frac{r_{1,P2}}{c_{3,M1}}\right)\left(\frac{\widehat{\Pi}_1(\mathbf{A}_{25}) - \Delta_{1,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \\ &-\left(\frac{r_{2,P1}}{c_{3,M1}}\right)\left(\frac{\widehat{\Pi}_2(\mathbf{A}_{25}) - \Delta_{2,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \\ &+\left(\frac{c_{4,M1}}{c_{3,M1}}\right)\left(\frac{\widehat{\Pi}_4(\mathbf{A}_{25}) + \Delta_{4,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \\ &+\left(\frac{EV(\mathbf{A}_{25})}{c_{3,M1}}\right)\begin{Bmatrix} \tau_{1,P2}\left(\frac{\widehat{\Pi}_1(\mathbf{A}_{25}) - \Delta_{1,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \\ +\tau_{2,P1}\left(\frac{\widehat{\Pi}_2(\mathbf{A}_{25}) - \Delta_{2,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \\ +\tau_{3,P1} + \tau_{4,M1}\left(\frac{\widehat{\Pi}_4(\mathbf{A}_{25}) + \Delta_{4,3}^{M1,P1}}{\widehat{\Pi}_3(\mathbf{A}_{25})}\right) \end{Bmatrix}.\end{aligned} \tag{6}$$

If $\gamma_3^{M1,P1} \leqslant 0$, then $a_3^* = P1$. However, if $\gamma_3^{M1,P1} > 0$, the optimal decision in state 3 can be determined by comparing $\gamma_3^{M1,P1}$ with $\frac{r_{3,P1}}{c_{3,M1}}$. (a) If $\frac{r_{3,P1}}{c_{3,M1}} > \gamma_3^{M1,P1}$, then $a_3^* = P1$ because $EV(\mathbf{A}_9 = [P2, P1, P1, M1]) > EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$; (b) If $\frac{r_{3,P1}}{c_{3,M1}} < \gamma_3^{M1,P1}$, then $a_3^* = M1$ because $EV(\mathbf{A}_9 = [P2, P1, P1, M1]) < EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$; and (c) If $\frac{r_{3,P1}}{c_{3,M1}} = \gamma_3^{M1,P1}$, then the firm is indifferent between production and maintenance actions in state 3 because $EV(\mathbf{A}_9 = [P2, P1, P1, M1]) = EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$.

There is an economic interpretation of the critical ratio in (6), which corresponds to the ratio of the profit that can be earned by producing in state 3 relative to the maintenance cost. The value of (6) tells the decision maker the least amount of money that she needs to earn in order to justify manufacturing over maintenance in a deteriorated intermediate state. Thus, the critical ratio $\gamma_3^{M1,P1}$ can be interpreted as the *reservation price* for the manufacturing option. Because $r_{i,a_i \in \mathbf{P}} > 0$ and $c_{i,M1} > 0$ for all i, a critical ratio value that is less than zero implies that the firm benefits more by the manufacturing option than the maintenance alternative. The value of the critical ratio $\gamma_3^{M1,P1}$ increases with: (i) lower values of $r_{1,P2}$ and $r_{2,P1}$, i.e., the profit earned from production in states 1 and 2; (ii) higher values of $c_{4,M1}$, the cost of the maintenance action in state 4; (iii) higher values of $\Delta_{1,3}^{M1,P1}$, $\Delta_{2,3}^{M1,P1}$ and $\Delta_{4,3}^{M1,P1}$, i.e., the change in the numerator terms in states 1, 2, and 4; (iv) higher values of expected processing times for production actions: $\tau_{1,P2}$, $\tau_{2,P1}$ and $\tau_{3,P1}$, and $\tau_{4,M1}$, the expected processing time of the maintenance action $M1$ in state 4; and (v) higher values of $EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$, the expected value generated from policy $\mathbf{A}_{25}$ featuring the maintenance action in state 3. All five of these conditions imply that the firm needs to earn a higher profit in state 3 in order to justify manufacturing of $P1$ rather than employing the maintenance action $M1$.

3.2. The production choice in an intermediate state

This section analyzes the scenario when the decision in the deteriorated state is restricted to manufacturing $P1$ or $P2$. The following two policies can be used in order to develop the critical ratio for the production choice in state 3: $\mathbf{A}_9 = [P2, P1, P1, M1]$ and $\mathbf{A}_{11} = [P2, P1, P2, M1]$. It was argued earlier that the process is more likely to deteriorate from state 3 to state 4 when product $P2$ is manufactured (rather than product $P1$); thus, $p_{33}^{P2} < p_{33}^{P1}$ and $p_{34}^{P2} > p_{34}^{P1}$. From (2), the firm has $\widehat{\Pi}_1(\mathbf{A}_9) = p_{41}^{M1}(1-p_{22}^{P1})(1-p_{33}^{P1}) < \widehat{\Pi}_1(\mathbf{A}_{11}) = p_{41}^{M1}(1-p_{22}^{P1})(1-p_{33}^{P2})$ because $(1-p_{33}^{P1}) < (1-p_{33}^{P2})$. For the same reason, from (3) and (5), it can be seen that $\widehat{\Pi}_2(\mathbf{A}_9) < \widehat{\Pi}_2(\mathbf{A}_{11})$ and $\widehat{\Pi}_4(\mathbf{A}_9) < \widehat{\Pi}_4(\mathbf{A}_{11})$. These observations imply that the numerators of the steady-state probabilities of states 1, 2 and 4 are greater in policy $\mathbf{A}_{11}$. As a result, the steady-state probability of state 3 is smaller in policy $\mathbf{A}_{11}$. Let us define $\delta_3^{P1,P2} = \frac{p_{ij}^{P2}}{p_{ij}^{P1}}$ as the ratio of the deterioration probabilities from producing the high-end product $P2$ and the low-end product $P1$ for all $1 \leqslant i < j \leqslant N$. From (1), $\delta_3^{P1,P2} > 1$ and has a finite value. Then, the relationship between numerator terms can be expressed as follows: $\widehat{\Pi}_1(\mathbf{A}_{11}) = \widehat{\Pi}_1(\mathbf{A}_9) \times \delta_3^{P1,P2}$; $\widehat{\Pi}_2(\mathbf{A}_{11}) = \widehat{\Pi}_2(\mathbf{A}_9) \times \delta_3^{P1,P2}$; $\widehat{\Pi}_3(\mathbf{A}_{11}) = \widehat{\Pi}_3(\mathbf{A}_9)$; and $\widehat{\Pi}_4(\mathbf{A}_{11}) = \widehat{\Pi}_3(\mathbf{A}_9) \times \delta_3^{P1,P2}$. The firm can now develop a critical ratio of revenues that determines the production choice in the intermediate state.

Proposition 2. *There exists a critical ratio that determines the manufacturing preference in state 3:*

$$\alpha_3^{P1,P2} = \delta_3^{P1,P2} + EV(\mathbf{A}_9 = [P2, P1, P1, M1])\left(\frac{\tau_{3,P2} - \tau_{3,P1}\delta_3^{P1,P2}}{r_{3,P1}}\right). \quad (7)$$

If $\alpha_3^{P1,P2} \leqslant 1$, then $a_3^* = P2$. However, if $\alpha_3^{P1,P2} > 1$, then the optimal production decision in state 3 can be determined by comparing $\alpha_3^{P1,P2}$ with $\frac{r_{3,P2}}{r_{3,P1}}$. (a) If $\frac{r_{3,P2}}{r_{3,P1}} > \alpha_3^{P1,P2}$, then $a_3^* = P2$ because $EV(\mathbf{A}_{11} = [P2, P1, P2, M1]) > EV(\mathbf{A}_9 = [P2, P1, P1, M1])$; (b) If $\frac{r_{3,P2}}{r_{3,P1}} < \alpha_3^{P1,P2}$, then $a_3^* = P1$ because $EV(\mathbf{A}_{11} = [P2, P1, P2, M1]) < EV(\mathbf{A}_9 = [P2, P1, P1, M1])$; and (c) If $\frac{r_{3,P2}}{r_{3,P1}} = \alpha_3^{P1,P2}$, then the firm is indifferent between manufacturing $P1$ and $P2$ in state 3 because $EV(\mathbf{A}_{11} = [P2, P1, P2, M1]) = EV(\mathbf{A}_9 = [P2, P1, P1, M1])$.

The critical ratio $\alpha_3^{P1,P2}$ provides the decision maker with a *reservation price* corresponding to her manufacturing choices. The firm needs to make at least $\alpha_3^{P1,P2} \times r_{3,P1}$ in state 3 in order to justify manufacturing the high-end product $P2$ rather than the standard product $P1$. A value of $\alpha_3^{P1,P2}$ that is less than or equal to 1 implies that the firm benefits more by producing $P2$. The value of the critical ratio increases with: (i) higher values of $\tau_{3,P2}$, the expected processing time of manufacturing $P2$ in state 3; (ii) lower values of $\tau_{3,P1}$, the expected processing time of manufacturing $P1$ in state 3; and (iii) the expected value gained from the policy that features the production of $P1$. These three observations lead to a higher profit requirement for the firm to switch from manufacturing product $P1$ to product $P2$. It can be seen that when $\frac{\tau_{3,P2}}{\tau_{3,P1}} > \delta_3^{P1,P2}$, the firm has to earn more money than $\delta_3^{P1,P2} \times r_{3,P1}$ by producing $P2$ in order to switch from policy $\mathbf{A}_9 = [P2, P1, P1, M1]$to $\mathbf{A}_{11} = [P2, P1, P2, M1]$. Moreover, $\alpha_3^{P1,P2}$ is less than $\delta_3^{P1,P2}$ only when $1 < \frac{\tau_{3,P2}}{\tau_{3,P1}} < \delta_3^{P1,P2}$; otherwise the critical ratio is always larger than the change that takes place in the numerators of steady-state probabilities.

It is important to highlight that (7) generalizes the similar critical ratios developed in Kazaz and Sloan (2008). In that paper, the transition probabilities are defined as linearly proportional with the expected processing times. Our transition probabilities, however, are general as no assumption is made regarding their relationship with the expected processing times.

The critical ratio in (7) provides insight into monotone and non-monotone policies. When the firm's ratio of profits earned in state from producing $P2$ and $P1$ is greater than the critical ratio (corresponding to the case when $\frac{r_{3,P2}}{r_{3,P1}} > \alpha_3^{P1,P2}$), the firm's optimal policy is $\mathbf{A}_{11} = [P2, P1, P2, M1]$ with the production action $P2$ in state 3. Because the firm produces $P1$ with a lower profit in a better state ($i = 2$), this case implies that a non-monotone policy is preferred. Because $\delta_3^{P1,P2}$, $EV(\mathbf{A}_9 = [P2, P1, P1, M1])$, and $r_{3,P1}$ are positive, the value of the critical ratio decreases only when $\tau_{3,P2} - \tau_{3,P1}\delta_3^{P1,P2} < 0$. If the difference in $\tau_{3,P2}$ and $\tau_{3,P1}$ is small, the critical ratio decreases with larger values of $\delta_3^{P1,P2}$, increasing the possibility that the non-monotone policy $\mathbf{A}_{11} = [P2, P1, P2, M1]$ would be preferred. When $\tau_{3,P2} - \tau_{3,P1}\delta_3^{P1,P2} > 0$, on the other hand, the firm has a critical ratio greater than the ratio of deterioration probabilities, i.e., $\alpha_3^{P1,P2} > \delta_3^{P1,P2}$. In this case, the possibility that the firm would prefer the non-monotone policy $\mathbf{A}_{11} = [P2, P1, P2, M1]$ decreases as $\delta_3^{P1,P2}$ increases. Thus, the higher the difference in the expected processing times of $P2$ and $P1$, the more likely that the firm will follow a monotone policy. A detailed discussion on the conditions that lead to monotone and non-monotone policies is provided in Section 5 using a more general problem setting.

3.3. The maintenance choice in an intermediate state

We now present the firm's maintenance preference in the deteriorated intermediate state. The decision is restricted to performing maintenance actions $M1$ and $M2$. The following two policies are beneficial in developing the critical ratio for the maintenance choice in state 2: $\mathbf{A}_{25} = [P2, P1, M1, M1]$ and $\mathbf{A}_{27} = [P2, P1, M2, M1]$. As mentioned earlier, the process is more likely to improve from state 3 to states 1 and 2 when maintenance action $M2$ is performed; thus, $p_{33}^{M2} < p_{33}^{M1}$ and $p_{3i}^{M2} > p_{3i}^{M1}$ for $i = 1, 2$. It can be seen from (2)–(5) that the firm has $\widehat{\Pi}_1(\mathbf{A}_{25}) < \widehat{\Pi}_1(\mathbf{A}_{27})$, $\widehat{\Pi}_2(\mathbf{A}_{25}) < \widehat{\Pi}_2(\mathbf{A}_{27})$, $\widehat{\Pi}_3(\mathbf{A}_{25}) = \widehat{\Pi}_3(\mathbf{A}_{27})$ and $\widehat{\Pi}_4(\mathbf{A}_{25}) < \widehat{\Pi}_3(\mathbf{A}_{27})$. These observations imply that the numerators of the steady-state probabilities of states 1, 2 and 4 are greater in policy $\mathbf{A}_{27}$. Therefore, the steady-state probability of state 3 is smaller in policy $\mathbf{A}_{27}$. Let us define $\delta_3^{M1,M2} = \frac{p_{ij}^{M2}}{p_{ij}^{M1}}$ as the ratio of improvement probabilities from utilizing maintenance actions $M2$ and $M1$ for all $1 \leqslant j < i \leqslant N$. From (1), $\delta_3^{M1,M2} > 1$ and has a finite value. The relationship between the numerator terms can be expressed as follows: $\widehat{\Pi}_1(\mathbf{A}_{27}) = \widehat{\Pi}_1(\mathbf{A}_{25}) \times \delta_3^{M1,M2}$; $\widehat{\Pi}_2(\mathbf{A}_{27}) = \widehat{\Pi}_2(\mathbf{A}_{25}) \times \delta_3^{M1,M2}$; $\widehat{\Pi}_3(\mathbf{A}_{27}) = \widehat{\Pi}_3(\mathbf{A}_{25})$ and $\widehat{\Pi}_4(\mathbf{A}_{27}) = \widehat{\Pi}_4(\mathbf{A}_{25}) \times \delta_3^{M1,M2}$. Using these relationships, the firm can develop another critical ratio in order to determine the maintenance choice in the intermediate state.

Proposition 3. *There exists a critical ratio that determines the maintenance preference in state 3:*

$$\lambda_3^{M1,M2} = \delta_3^{M1,M2} + EV(\mathbf{A}_{25} = [P2, P1, M1, M1])\left(\frac{\tau_{3,M1}\delta_3^{M1,M2} - \tau_{3,M2}}{c_{3,M1}}\right). \quad (8)$$

If $\lambda_3^{M1,M2} \leqslant 1$, then $a_3^* = M1$. However, if $\lambda_3^{M1,M2} > 1$, then the optimal maintenance decision in state 3 can be determined by comparing $\lambda_3^{M1,M2}$ with $\frac{c_{3,M2}}{c_{3,M1}}$. (a) If $\frac{c_{3,M2}}{c_{3,M1}} > \lambda_3^{M1,M2}$, then $a_3^* = M1$ because $EV(\mathbf{A}_{27} = [P2, P1, M2, M1]) < EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$; (b) If $\frac{c_{3,M2}}{c_{3,M1}} < \lambda_3^{M1,M2}$, then $a_3^* = M2$ because $EV(\mathbf{A}_{27} = [P2, P1, M2, M1]) > EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$; and (c) If $\frac{c_{3,M2}}{c_{3,M1}} = \lambda_3^{M1,M2}$, then the firm is indifferent between maintenance actions $M1$ and $M2$ in state 2 because $EV(\mathbf{A}_{27} = [P2, P1, M2, M1]) = EV(\mathbf{A}_{25} = [P2, P1, M1, M1])$.

A similar economic interpretation can be made for the critical ratio $\lambda_3^{M1,M2}$ representing the ratio of maintenance expenses between a major and a minor maintenance. It provides the decision maker with the *maximum* amount of money to be spent in order to justify using the major maintenance action $M2$ over the standard action $M1$. Specifically, $\lambda_3^{M1,M2} \times c_{3,M1}$ is the highest amount of money that the firm should be willing to pay for a major maintenance action in a deteriorated intermediate state. If the maintenance cost of $M2$ is lower than this amount, then the firm prefers to utilize a major maintenance; otherwise, it should continue to use the standard (or minor) maintenance action. The value of the critical ratio $\lambda_3^{M1,M2}$ increases with: (i) lower values of $\tau_{3,M2}$, the expected processing time of the major maintenance action $M2$ in state 3; (ii) higher values of $\tau_{3,M1}$, the expected processing time of the minor maintenance action $M1$ in state 3; and (iii) the expected value gained from the policy that features the minor maintenance action $M1$ in state 3. These observations lead to a higher upper bound, increasing the likelihood of employing the major maintenance action $M2$ in the optimal decision. It can be seen that when $\frac{\tau_{3,M2}}{\tau_{3,M1}} > \delta_3^{M1,M2}$, the firm has to spend less money than $\lambda_3^{M1,M2} \times c_{3,M1}$ for the maintenance action $M2$ in order to switch from policy $\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$] to $\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$]. Moreover, $\lambda_3^{M1,M2}$ is less than $\delta_3^{M1,M2}$ only when $\frac{\tau_{3,M2}}{\tau_{3,M1}} > \delta_3^{M1,M2} > 1$; otherwise the critical ratio is always larger than the change that takes place in the numerators of steady-state probabilities.

The critical ratio in (8) provides insight into the maintenance-related monotone and non-monotone policies. When the firm's ratio of maintenance costs from the major maintenance $M2$ and the minor maintenance $M1$ is less than the critical ratio $\lambda_3^{M1,M2}$, the firm's optimal policy is $\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$] with the maintenance action $M2$ in state 3. Because the firm utilizes $M1$ with a lower expense in a worse state, this case implies that a non-monotone policy is preferred. Because $\delta_3^{M1,M2}$, $EV(\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$]), and $c_{3,M1}$ are positive, the value of the critical ratio increases only when $\tau_{3,M1}\delta_3^{M1,M2} - \tau_{3,M2} > 0$. In other words, if the difference in $\tau_{3,M2}$ and $\tau_{3,M1}$ is small, the critical ratio increases with larger values of $\delta_3^{M1,M2}$, making it easier for the firm to prefer the non-monotone policy $\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$]. Moreover, when $\tau_{3,M1}\delta_3^{M1,M2} - \tau_{3,M2} < 0$, the firm has a critical ratio less than the ratio of improvement probabilities, i.e., $\lambda_3^{M1,M2} < \delta_3^{M1,M2}$. In this case, the possibility that the firm would prefer the non-monotone policy $\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$] becomes increasingly difficult. Thus, the higher the difference in the expected processing times of $M1$ and $M2$, the more likely that the firm will follow a monotone policy. A detailed discussion regarding the conditions for monotonicity is provided in Section 5.

3.4. Combining the three critical ratios

This section shows how the critical ratios can be combined in order to determine the best policy among the four candidate policies: $\mathbf{A}_9$ = [$P2$, $P1$, $P1$, $M1$], $\mathbf{A}_{11}$ = [$P2$, $P1$, $P2$, $M1$], $\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$], and $\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$]. Recall that the critical ratio $\gamma_3^{M1,P1}$ enables the firm to choose between the production and maintenance options, $\alpha_3^{P1,P2}$ provides the best production alternative, and $\lambda_3^{M1,M2}$ reveals the best maintenance action. They help the decision maker to determine the best policy.

Proposition 4. *(a) The best policy is* $\mathbf{A}_9$ *= [P2, P1, P1, M1] when* $r_{3,P1} \geqslant \left\{ r_{3,P2}\frac{1}{\alpha_3^{P1,P2}}, c_{3,M1}\gamma_3^{M1,P1}, c_{3,M2}\frac{\gamma_3^{M1,P1}}{\lambda_3^{M1,M2}} \right\}$*; (b) The best policy is* $\mathbf{A}_{11}$ *= [P2, P1, P2, M1] when* $r_{3,P2} \geqslant \left\{ r_{3,P1}\alpha_3^{P1,P2}, c_{3,M1}\gamma_3^{M1,P1}\alpha_3^{P1,P2}, c_{3,M2}\frac{\gamma_3^{M1,P1}}{\lambda_3^{M1,M2}}\alpha_3^{P1,P2} \right\}$*; (c) The best policy is* $\mathbf{A}_{25}$ *= [P2, P1, M1, M1] when* $c_{3,M1} \geqslant \left\{ r_{3,P1}\frac{1}{\gamma_3^{M1,P1}}, r_{3,P2}\frac{1}{\gamma_3^{M1,P1}\alpha_3^{P1,P2}} \right\}$ *and* $c_{3,M1} \leqslant c_{3,M2}\frac{1}{\lambda_3^{M1,M2}}$*; (d) The best policy is* $\mathbf{A}_{27}$ *= [P2, P1, M2, M1] when* $c_{3,M2} \geqslant \left\{ r_{3,P1}\frac{\lambda_3^{M1,M2}}{\gamma_3^{M1,P1}}, r_{3,P2}\frac{\lambda_3^{M1,M2}}{\gamma_3^{M1,P1}\alpha_3^{P1,P2}} \right\}$ *and* $c_{3,M2} \leqslant c_{3,M1}\lambda_3^{M1,M2}$.

Proposition 4 enables the firm to determine the optimal choice in state 3 when the decisions in other states are restricted to $a_1 = P2$, $a_2 = P1$ and $a_4 = M1$. However, the firm has a production choice in state 1 ($a_1 \in \mathbf{P}$), the same four choices in state 2 ($a_2 \in \mathbf{P} \cup \mathbf{M}$) and a maintenance choice in state 4 ($a_4 \in \mathbf{M}$). The same set of critical ratios can be developed for other states. It is necessary to develop $\gamma_2^{M1,P1}$ for state 2, $\alpha_1^{P1,P2}$ and $\alpha_2^{P1,P2}$ for states 1 and 2, and $\lambda_2^{M1,M2}$ and $\lambda_4^{M1,M2}$ for states 2 and 4 in order to determine the optimal policy among the previously reported 64 policies. Section 5 presents a comprehensive review of the generalized forms of the critical ratios using an arbitrary number of states.

4. Incorporating minimum and maximum production requirements

An important issue for semiconductor manufacturers is to comply with market requirements, which corresponds to the firm's commitment to producing an expected amount of each of its products. The comprehensive list of policies in a four-state problem provided in Table 1 includes many pure product policies such as $\mathbf{A}_1$ = [$P1$, $P1$, $P1$, $M1$], where only product $P1$ is manufactured. When optimal, policy $\mathbf{A}_1$ implies that product $P2$ should not be manufactured. However, it is likely that the firm will be operating under demand constraints, requiring that it manufacture both products. Incorporating production requirements has a significant impact both on the optimal policy choice and the critical ratios available for comparison.

Under a policy $\mathbf{A}$, the expected production quantities for each product are defined as $Y_{P1}(\mathbf{A}) = \sum_{i=1}^{N} y_{ia_i}\, \Pi_i(\mathbf{A}) 1_{a_i=P_1} / \sum_{i=1}^{N} \tau_{i,a_i} \Pi_i(\mathbf{A})$ and $Y_{P2}(\mathbf{A}) = \sum_{i=1}^{N} y_{ia_i} \Pi_i(\mathbf{A}) 1_{a_i=P_2} / \sum_{i=1}^{N} \tau_{i,a_i} \Pi_i(\mathbf{A})$. Based on obligations to downstream electronics manufacturers, for example, the firm might enforce a minimum on the expected production quantity for $P1$ and $P2$, defined as MPR_{P1} and MPR_{P2}, respectively, through the following constraints:

$$Y_{P1}(\mathbf{A}) \geqslant MPR_{P1} \quad \text{and} \quad Y_{P2}(\mathbf{A}) \geqslant MPR_{P2}. \tag{9}$$

The immediate consequence of non-negative MPR_{P1} and MPR_{P2} constraints as in (9) is that it reduces the number of potentially optimal policies from 64 to 28 as shown in Table 2. It can be seen that stronger constraints on the minimum production requirements reduces this number even further.

Similarly, the administration might enforce a maximum production amount for its products, defined as XPR_{P1} and XPR_{P2}, respectively, as in the following constraints:

$$Y_{P1}(\mathbf{A}) \leqslant XPR_{P1} \quad \text{and} \quad Y_{P2}(\mathbf{A}) \leqslant XPR_{P2}. \tag{10}$$

It is important to note that semiconductor manufacturers generally enforce minimum production requirements in their production plans, but rarely introduce a maximum production requirement for their high-end products. This is because the firm can always downwardly substitute its high-end product in order to satisfy the unmet demand in its low-end product. Reflecting the operating environment at semiconductor manufacturers, even though we present the influence of minimum and maximum production requirements, we focus on the high-end product $P2$ in the presentation of minimum production requirements, and on the low-end product $P1$ in the presentation of maximum production limitations. Our proposed solution approach is rather general and applicable in alternative production environments, and therefore, we provide a comprehensive review of the implications of such minimum and maximum production requirement constraints on the optimal policy.

4.1. Impact of minimum production requirements

We first present how the firm can increase its throughput when the minimum production requirements for the high-end and low-

Table 1
Comprehensive list of policies in a four-state problem.

Group 1: [**P**, **P**, **P**, **M**]	Group 2: [**P**, **P**, **M**, **M**]	Group 3: [**P**, **M**, **P**, **M**]	Group 4: [**P**, **M**, **M**, **M**]
$\mathbf{A}_1$ = [$P1$, $P1$, $P1$, $M1$]	$\mathbf{A}_{17}$ = [$P1$, $P1$, $M1$, $M1$]	$\mathbf{A}_{33}$ = [$P1$, $M1$, $P1$, $M1$]	$\mathbf{A}_{49}$ = [$P1$, $M1$, $M1$, $M1$]
$\mathbf{A}_2$ = [$P1$, $P1$, $P1$, $M2$]	$\mathbf{A}_{18}$ = [$P1$, $P1$, $M1$, $M2$]	$\mathbf{A}_{34}$ = [$P1$, $M1$, $P1$, $M2$]	$\mathbf{A}_{50}$ = [$P1$, $M1$, $M1$, $M2$]
$\mathbf{A}_3$ = [$P1$, $P1$, $P2$, $M1$]	$\mathbf{A}_{19}$ = [$P1$, $P1$, $M2$, $M1$]	$\mathbf{A}_{35}$ = [$P1$, $M1$, $P2$, $M1$]	$\mathbf{A}_{51}$ = [$P1$, $M1$, $M2$, $M1$]
$\mathbf{A}_4$ = [$P1$, $P1$, $P2$, $M2$]	$\mathbf{A}_{20}$ = [$P1$, $P1$, $M2$, $M2$]	$\mathbf{A}_{36}$ = [$P1$, $M1$, $P2$, $M2$]	$\mathbf{A}_{52}$ = [$P1$, $M1$, $M2$, $M2$]
$\mathbf{A}_5$ = [$P1$, $P2$, $P1$, $M1$]	$\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$]	$\mathbf{A}_{37}$ = [$P1$, $M2$, $P1$, $M1$]	$\mathbf{A}_{53}$ = [$P1$, $M2$, $M1$, $M1$]
$\mathbf{A}_6$ = [$P1$, $P2$, $P1$, $M2$]	$\mathbf{A}_{22}$ = [$P1$, $P2$, $M1$, $M2$]	$\mathbf{A}_{38}$ = [$P1$, $M2$, $P1$, $M2$]	$\mathbf{A}_{54}$ = [$P1$, $M2$, $M1$, $M2$]
$\mathbf{A}_7$ = [$P1$, $P2$, $P2$, $M1$]	$\mathbf{A}_{23}$ = [$P1$, $P2$, $M2$, $M1$]	$\mathbf{A}_{39}$ = [$P1$, $M2$, $P2$, $M1$]	$\mathbf{A}_{55}$ = [$P1$, $M2$, $M2$, $M1$]
$\mathbf{A}_8$ = [$P1$, $P2$, $P2$, $M2$]	$\mathbf{A}_{24}$ = [$P1$, $P2$, $M2$, $M2$]	$\mathbf{A}_{40}$ = [$P1$, $M2$, $P2$, $M2$]	$\mathbf{A}_{56}$ = [$P1$, $M2$, $M2$, $M2$]
$\mathbf{A}_9$ = [$P2$, $P1$, $P1$, $M1$]	$\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$]	$\mathbf{A}_{41}$ = [$P2$, $M1$, $P1$, $M1$]	$\mathbf{A}_{57}$ = [$P2$, $M1$, $M1$, $M1$]
$\mathbf{A}_{10}$ = [$P2$, $P1$, $P1$, $M2$]	$\mathbf{A}_{26}$ = [$P2$, $P1$, $M1$, $M2$]	$\mathbf{A}_{42}$ = [$P2$, $M1$, $P1$, $M2$]	$\mathbf{A}_{58}$ = [$P2$, $M1$, $M1$, $M2$]
$\mathbf{A}_{11}$ = [$P2$, $P1$, $P2$, $M1$]	$\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$]	$\mathbf{A}_{43}$ = [$P2$, $M1$, $P2$, $M1$]	$\mathbf{A}_{59}$ = [$P2$, $M1$, $M2$, $M1$]
$\mathbf{A}_{12}$ = [$P2$, $P1$, $P2$, $M2$]	$\mathbf{A}_{28}$ = [$P2$, $P1$, $M2$, $M2$]	$\mathbf{A}_{44}$ = [$P2$, $M1$, $P2$, $M2$]	$\mathbf{A}_{60}$ = [$P2$, $M1$, $M2$, $M2$]
$\mathbf{A}_{13}$ = [$P2$, $P2$, $P1$, $M1$]	$\mathbf{A}_{29}$ = [$P2$, $P2$, $M1$, $M1$]	$\mathbf{A}_{45}$ = [$P2$, $M2$, $P1$, $M1$]	$\mathbf{A}_{61}$ = [$P2$, $M2$, $M1$, $M1$]
$\mathbf{A}_{14}$ = [$P2$, $P2$, $P1$, $M2$]	$\mathbf{A}_{30}$ = [$P2$, $P2$, $M1$, $M2$]	$\mathbf{A}_{46}$ = [$P2$, $M2$, $P1$, $M2$]	$\mathbf{A}_{62}$ = [$P2$, $M2$, $M1$, $M2$]
$\mathbf{A}_{15}$ = [$P2$, $P2$, $P2$, $M1$]	$\mathbf{A}_{31}$ = [$P2$, $P2$, $M2$, $M1$]	$\mathbf{A}_{47}$ = [$P2$, $M2$, $P2$, $M1$]	$\mathbf{A}_{63}$ = [$P2$, $M2$, $M2$, $M1$]
$\mathbf{A}_{16}$ = [$P2$, $P2$, $P2$, $M2$]	$\mathbf{A}_{32}$ = [$P2$, $P2$, $M2$, $M2$]	$\mathbf{A}_{48}$ = [$P2$, $M2$, $P2$, $M2$]	$\mathbf{A}_{64}$ = [$P2$, $M2$, $M2$, $M2$]

Table 2
List of policies that feature the manufacturing of both products.

Group 1: [P, P, P, M]	Group 2: [P, P, M, M]	Group 3: [P, M, P, M]
$\mathbf{A}_3$ = [$P1$, $P1$, $P2$, $M1$]	$\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$]	$\mathbf{A}_{35}$ = [$P1$, $M1$, $P2$, $M1$]
$\mathbf{A}_4$ = [$P1$, $P1$, $P2$, $M2$]	$\mathbf{A}_{22}$ = [$P1$, $P2$, $M1$, $M2$]	$\mathbf{A}_{36}$ = [$P1$, $M1$, $P2$, $M2$]
$\mathbf{A}_5$ = [$P1$, $P2$, $P1$, $M1$]	$\mathbf{A}_{23}$ = [$P1$, $P2$, $M2$, $M1$]	$\mathbf{A}_{39}$ = [$P1$, $M2$, $P2$, $M1$]
$\mathbf{A}_6$ = [$P1$, $P2$, $P1$, $M2$]	$\mathbf{A}_{24}$ = [$P1$, $P2$, $M2$, $M2$]	$\mathbf{A}_{40}$ = [$P1$, $M2$, $P2$, $M2$]
$\mathbf{A}_7$ = [$P1$, $P2$, $P2$, $M1$]	$\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$]	$\mathbf{A}_{41}$ = [$P2$, $M1$, $P1$, $M1$]
$\mathbf{A}_8$ = [$P1$, $P2$, $P2$, $M2$]	$\mathbf{A}_{26}$ = [$P2$, $P1$, $M1$, $M2$]	$\mathbf{A}_{42}$ = [$P2$, $M1$, $P1$, $M2$]
$\mathbf{A}_9$ = [$P2$, $P1$, $P1$, $M1$]	$\mathbf{A}_{27}$ = [$P2$, $P1$, $M2$, $M1$]	$\mathbf{A}_{45}$ = [$P2$, $M2$, $P1$, $M1$]
$\mathbf{A}_{10}$ = [$P2$, $P1$, $P1$, $M2$]	$\mathbf{A}_{28}$ = [$P2$, $P1$, $M2$, $M2$]	$\mathbf{A}_{46}$ = [$P2$, $M2$, $P1$, $M2$]
$\mathbf{A}_{11}$ = [$P2$, $P1$, $P2$, $M1$]		
$\mathbf{A}_{12}$ = [$P2$, $P1$, $P2$, $M2$]		
$\mathbf{A}_{13}$ = [$P2$, $P2$, $P1$, $M1$]		
$\mathbf{A}_{14}$ = [$P2$, $P2$, $P1$, $M2$]		

end products do not satisfy MPR_{P1} and MPR_{P2} constraints in (9). We present the analysis for the high-end product, and similar conditions can be developed for the low-end product. Let us consider the event that the optimal policy violates the MPR_{P2} in (9) and that $r_{3,P2}/c_{3,M1} < \gamma_3^{M1,P1}\alpha_3^{P1,P2}$, implying that the optimal action in state 3 from an economic perspective is $M1$ (by Proposition 4). Note that the firm performs maintenance in the worst state, so the unichain property of the SMDP is preserved; as a result, production takes place only in states 1 and 2, corresponding to the policies in Group 2 in Table 2.

Because the expected yield is smaller in state 2 for both products, we consider policy $\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$] as the reference policy but assume that it violates the MPR_{P2} constraint in (9). When this is the case, we show that the firm can increase its $P2$ throughput in three different ways. First, the firm can switch from maintenance to manufacturing its high-end product in a deteriorated state; in particular, it might choose to manufacture $P2$ instead of performing maintenance in state 3. The comparison of the expected production from policies $\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$] and $\mathbf{A}_7$ = [$P1$, $P2$, $P2$, $M1$], provides the conditions for increasing the throughput. Second, switching to a major maintenance from a minor maintenance action can increase throughput in better states, i.e., states 1 and 2. In this case, the comparison of policies $\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$] and $\mathbf{A}_{23}$ = [$P1$, $P2$, $M2$, $M1$] shows the conditions to increase the expected production.

Proposition 5. *The firm can increase its expected production of P2 by: (a) switching from maintenance to production in an intermediate state, e.g., state 3, when* $y_{3,P2} - y_{2,P2}(\Delta_{2,3}^{M1,P2}/\widehat{\Pi}_3(\mathbf{A}_{21})) > Y_{P2}(\mathbf{A}_{21})$

$$\left\{\begin{array}{c} -\tau_{1,P1}\left(\Delta_{1,3}^{M1,P2}/\widehat{\Pi}_3(\mathbf{A}_{21})\right) - \tau_{2,P2}\left(\Delta_{2,3}^{M1,P2}/\widehat{\Pi}_3(\mathbf{A}_{21})\right) \\ +(\tau_{3,P2} - \tau_{3,M1}) + \tau_{4,M1}\left(\Delta_{4,3}^{M1,P2}/\widehat{\Pi}_3(\mathbf{A}_{21})\right) \end{array}\right\}$$

; (b) applying major maintenance, rather than minor maintenance, in an intermediate state when $\tau_{3,M2} - \tau_{3,M1}\delta_3^{M1,M2} < 0$.

A third alternative to increasing the expected output of the high-end product involves performing maintenance in an earlier (better) state in order to increase the frequency of manufacturing in the best states. This requires swapping of production and maintenance in a better state. This can be seen in the comparison of the expected production amounts from policies $\mathbf{A}_{25}$ = [$P2$, $P1$, $M1$, $M1$] and $\mathbf{A}_{41}$ = [$P2$, $M1$, $P1$, $M1$]. Note that there is a double switch, from $P1$ to $M1$ in state 2 and from $M1$ to $P1$ in state 3, in this comparison. The following proposition shows the necessary and sufficient condition for this action.

Proposition 6. *The firm can increase the throughput of a product, e.g., P2, by performing maintenance in a better state when*

$$y_{1,P2} > \frac{Y_{P2}(\mathbf{A}_{25})}{\left(\Delta_{1,2}^{M1,P1} - \Delta_{1,3}^{M1,P1}\right)}\left\{\begin{array}{c} \tau_{1,P2}\left(\Delta_{1,2}^{M1,P1} - \Delta_{1,3}^{M1,P1}\right) \\ +(\tau_{2,M1} - \tau_{2,P1})\widehat{\Pi}_2(\mathbf{A}_{25}) - \tau_{2,M1}\Delta_{2,3}^{M1,P1} \\ +(\tau_{3,P1} - \tau_{3,M1})\widehat{\Pi}_3(\mathbf{A}_{25}) + \tau_{3,P1}\Delta_{3,2}^{M1,P1} \\ +\tau_{4,M1}\left(\Delta_{4,3}^{M1,P1} - \Delta_{4,2}^{M1,P1}\right) \end{array}\right\}.$$

Three conclusions can be made from the above two propositions. Part (a) of Proposition 5 shows that the firm can increase its expected production of $P2$ when the yield in state 3 is relatively close to that of state 2 and the changes in steady-state probabilities do not increase the adjusted expected total processing time (right hand side of the condition). This condition is not satisfied when the firm is spending too much time performing maintenance, resulting in lower throughput. Part (b) of the same proposition proves that the firm can increase the throughput of a product by switching from minor maintenance to major maintenance. This occurs when the expected processing time of $M2$ is smaller than that of $M1$ multiplied by the increase in improvement probabilities. Proposition 6 shows that the firm can also increase the throughput of a product manufactured in better states by performing its maintenance in an earlier state. This requires that adjustment in the total expected processing time is not significant. Considering the results in Propositions 5 and 6, it can be concluded that the *frequency* and *timing* of maintenance can play a strategic role in increasing the throughput of a product. This can be accomplished by either performing major maintenance or performing maintenance before the process becomes highly deteriorated. Finally, it should be stated that when the firm needs to increase the expected throughput of its low-end product $P1$, it can develop similar conditions to those presented in Propositions 5 and 6. These conditions are omitted in the manuscript for two reasons: (1) in order to focus on the development of structural properties, and (2) to reflect the operating environment for a semiconductor manufacturer who would enforce a minimum production requirement on its high-end product, or alternatively, a maximum production amount on its low-end product.

A comparison of the expected production quantities reveals that policies in Group 3 result in the lowest expected production levels for each product. Specifically, $\mathbf{A}_{45}$ = [$P2$, $M2$, $P1$, $M1$] and $\mathbf{A}_{39}$ = [$P1$, $M2$, $P2$, $M1$] provide the smallest $Y_{P1}(\mathbf{A})$ and $Y_{P2}(\mathbf{A})$ for products $P1$ and $P2$, respectively. Consider the event that $Y_{P1}(\mathbf{A}_{45})$ does not satisfy (9), making $\mathbf{A}_{45}$ infeasible. Then, the next three policies with the smallest expected production quantities are $\mathbf{A}_{46}$ = [$P2$, $M2$, $P1$, $M2$], $\mathbf{A}_{41}$ = [$P2$, $M1$, $P1$, $M1$], and $\mathbf{A}_{42}$ = [$P2$, $M1$, $P1$, $M2$]. If these three policies also fail to satisfy (9), then the decision maker has to consider producing $P1$ at the latest in state 2 or better (state 1), eliminating these four policies from further consideration. The following proposition shows that the comparison of the maximum $Y_{P1}(\mathbf{A})$ and $Y_{P2}(\mathbf{A})$ from policies that feature manufacturing of each product in a single state with MPR_{P1} and MPR_{P2} leads to an effective set of structural properties.

Proposition 7. *Under the conditions established in Propositions 5 and 6 for increasing throughput (a) If $MPR_{P1} > Y_{P1}(\mathbf{A}_{42})$ and $MPR_{P2} > Y_{P2}(\mathbf{A}_{36})$, then no policy in Group 3 can be optimal. Thus, the number of potentially optimal policies reduces to 20 and the critical ratios $\gamma_2^{M1,P1}$ and $\lambda_2^{M1,M2}$ can be eliminated from the problem. (b) If $MPR_{P1} > Y_{P1}(\mathbf{A}_{24})$ or $MPR_{P2} > Y_{P2}(\mathbf{A}_{28})$, then no policy in Group 2 can be optimal. Thus, the number of potentially optimal policies reduces to 20 and the critical ratios $\gamma_3^{M1,P1}$ and $\lambda_3^{M1,M2}$ can be eliminated from the problem. (c) If $MPR_{P1} > Y_{P1}(\mathbf{A}_{40})$ or $MPR_{P2} > Y_{P2}(\mathbf{A}_{46})$, then no policy in Groups 2 and 3 can be optimal. These conditions reduce the number of potentially optimal policies to 12 and the critical ratios $\gamma_2^{M1,P1}$, $\lambda_2^{M1,M2}$, $\gamma_3^{M1,P1}$ and $\lambda_3^{M1,M2}$ can be eliminated from the problem. Moreover, (d) when $MPR_{P1} > Y_{P1}(\mathbf{A}_{40})$ policies $\mathbf{A}_3$ through $\mathbf{A}_8$ in Group 1 and when $MPR_{P2} > Y_{P2}(\mathbf{A}_{46})$ policies $\mathbf{A}_9$ through $\mathbf{A}_{14}$ in Group 1 cannot be optimal, eliminating the critical ratios $\alpha_1^{P1,P2}$, $\gamma_2^{M1,P1}$, $\lambda_2^{M1,M2}$, $\gamma_3^{M1,P1}$ and $\lambda_3^{M1,M2}$ from the problem.*

The above proposition provides insight into the influence of the firm's minimum production requirements. First, it shows that stronger constraints reduce the number of potentially optimal policies, resulting in a smaller search for the optimal policy. Second, it highlights the relationship between the minimum production requirements and the three sets of critical ratios established in Sections 3.1, 3.2, 3.3. Stronger production requirements in (9) result in a smaller set of necessary critical ratios in determining the optimal policy. Example 1 in the Appendix demonstrates the impact of the minimum production requirements on the optimal policy choice.

4.2. Impact of maximum production constraints

We next describe the influence of constraints that limit the maximum amount of production for a particular product. A semiconductor manufacturer may enforce this condition on its low-end product, and therefore, we present our structural results by focusing on $P1$; similar results exist for the high-end product $P2$.

Let us consider the event that the optimal policy violates the XPR_{P1} constraint in (10) and that $r_{3,P1} > \left\{\gamma_3^{M1,P2} c_{3,M1}, r_3^{P1,P2}/\alpha_3^{P1,P2}\right\}$, implying that the optimal action in state 3 from an economic perspective is $P1$ (by Proposition 4). To develop our conditions, we consider $\mathbf{A}_5$ = [$P1$, $P2$, $P1$, $M1$] as the reference policy, but assume that it violates the maximum production amount XPR_{P1} in constraint (10). When this is the case, we show that the firm can reduce its $P1$ throughput in two different ways. First, the firm can switch from manufacturing its low-end product to maintenance in a deteriorated state; in particular, it might choose to maintain (with action $M1$) rather than manufacturing $P1$ in state 3. The comparison of the expected production amounts in policies $\mathbf{A}_5$ = [$P1$, $P2$, $P1$, $M1$] and $\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$] provides the conditions for decreasing the throughput. Second, switching from a major maintenance to a minor maintenance action can also decrease the throughput in better states, i.e., states 1 and 2. This is exemplified by comparing $\mathbf{A}_{23}$ = [$P1$, $P2$, $M2$, $M1$] and $\mathbf{A}_{21}$ = [$P1$, $P2$, $M1$, $M1$].

Proposition 8. *The firm can reduce its expected production of P1 by: (a) switching from production to maintenance in an intermediate state, e.g., state 3, when $y_{3,P1} - y_{2,P1}(\Delta_{2,3}^{M1,P1}/\widehat{\Pi}_3(\mathbf{A}_{21})) < Y_{P1}(\mathbf{A}_{21})\left\{\begin{matrix} -\tau_{1,P1}(\Delta_{1,3}^{M1,P1}/\widehat{\Pi}_3(\mathbf{A}_{21})) - \tau_{2,P2}(\Delta_{2,3}^{M1,P1}/\widehat{\Pi}_3(\mathbf{A}_{21})) \\ +(\tau_{3,P1} - \tau_{3,M1}) + \tau_{4,M1}(\Delta_{4,3}^{M1,P1}/\widehat{\Pi}_3(\mathbf{A}_{21})) \end{matrix}\right\}$; (b) applying minor maintenance rather than major maintenance in an intermediate state when $\tau_{3,M2} - \tau_{3,M1}\delta_3^{M1,M2} > 0$.*

Maximum production limitations can reduce the set of potentially optimal policies. Specifically, when the expected production amount from policies $\mathbf{A}_9$ = [$P2$, $P1$, $P1$, $M1$] and $\mathbf{A}_7$ = [$P1$, $P2$, $P2$, $M1$] yield higher values of $P1$ and $P2$, respectively, exceeding the maximum production amounts XPR_{P1} and XPR_{P2}, then no policy in Group 1 of Table 2 can be a viable alternative. Thus, the firm needs to switch from manufacturing in deteriorated states (e.g., state 3) to maintenance in order to reduce its expected yield. Moreover, when the expected yield from policies $\mathbf{A}_{35}$ = [$P1$, $M1$, $P2$, $M1$] and $\mathbf{A}_{41}$ = [$P2$, $M1$, $P1$, $M1$] exceed XPR_{P1} and XPR_{P2} limitations in constraints (10), then all policies in Group 3 of Table 2, and four other policies of Group 1 can be eliminated from the list of potentially optimal policies.

Proposition 9. *(a) If $XPR_{P1} > Y_{P1}(\mathbf{A}_9)$ and $XPR_{P2} > Y_{P2}(\mathbf{A}_7)$, then no policy in Group 1 can be optimal. Thus, the number of potentially optimal policies reduces to 16. (b) If $XPR_{P1} > Y_{P1}(\mathbf{A}_{35})$ and $XPR_{P2} > Y_{P2}(\mathbf{A}_{41})$, then policies in Group 3 and policies $\mathbf{A}_3$ = [P1, P1, P2, M1], $\mathbf{A}_4$ = [P1, P1, P2, M2], $\mathbf{A}_{13}$ = [P2, P2, P1, M1] and $\mathbf{A}_{14}$ = [P2, P2, P1, M2] in Group 1 cannot be optimal. Thus, the number of potentially optimal policies reduces to 16, and the critical ratios $\gamma_2^{M1,P1}$ and $\lambda_2^{M1,M2}$ can be eliminated from the problem.*

Note that the above proposition excludes the possibility of a double switch with postponed production of $P1$ until the equipment deteriorates to state 3 as is the case in policies in Group 3. This is because, as stated in Proposition 7, these policies can violate the minimum production requirement for $P1$. Thus, it can be concluded that incorporating minimum and maximum production requirements tend to push the optimal policy towards those presented in Group 2. Policies in Group 1 have the likelihood of violating the maximum production amount constraints in (10), and policies in Group 3 are likely not to satisfy the minimum production requirements enforced by constraints in (9).

It should be highlighted here that 12 of the 28 policies that manufacture both products in Table 1 are production-related monotone policies, i.e. as the state gets worse, the firm does not switch to a more profitable product: $\mathbf{A}_9$, $\mathbf{A}_{10}$, $\mathbf{A}_{13}$, $\mathbf{A}_{14}$, $\mathbf{A}_{25}$, $\mathbf{A}_{26}$, $\mathbf{A}_{27}$, $\mathbf{A}_{28}$, $\mathbf{A}_{41}$, $\mathbf{A}_{42}$, $\mathbf{A}_{45}$ and $\mathbf{A}_{46}$. However, five of these twelve policies violate this behavior from a maintenance perspective because the firm either switches to major maintenance or switches from production to maintenance as the state improves: $\mathbf{A}_{27}$, $\mathbf{A}_{41}$, $\mathbf{A}_{42}$, $\mathbf{A}_{45}$ and $\mathbf{A}_{46}$. As a result, there are *only* seven pure monotone policies that comply with production and maintenance action switches. Section 5 develops the general conditions for the optimality of production- and maintenance-related monotone and non-monotone policies.

Our goal in this paper is to present the structural properties of the problem and provide insight into the firm's decisions with the use of critical ratios. Therefore, we next focus on how these critical ratios change in problem settings with an arbitrary number of states.

5. Generalizing the critical ratios

The three sets of critical ratios developed in Section 3 can be generalized to a problem setting that features N states. Let us begin our discussion with the critical ratio that determines the firm's preference between the manufacturing and maintenance alternatives.

5.1. Critical ratios for switching between production and maintenance

We consider the policy $\mathbf{A}_n = [a_1, \ldots, a_N]$ with action $a_j = M1$ in state j and the firm needs to determine whether to switch its action to $a_j = P1$. To establish the general form of the critical ratio, it is necessary to highlight the following three changes in the numerator terms: (1) the numerator for the steady-state probability of state j remains the same, (2) the numerator term in states $i = 1, \ldots, j-1$ decreases with $\widehat{\Pi}_i(\mathbf{A}_n) - \Delta_{i,j}^{M1,P1}(\mathbf{A}_n)$, and (3) the numerator term in states $i = j+1, \ldots, N$ increases with $\widehat{\Pi}_i(\mathbf{A}_n) + \Delta_{i,j}^{M1,P1}(\mathbf{A}_n)$. As a result of these observations, the critical ratio corresponding to the choice between production and maintenance can be expressed as follows:

$$\begin{aligned}\gamma_j^{M1,P1} = &\sum_{i=1}^{j-1}\left[\left(-1_{a_i\in\mathbf{P}}\left(\frac{r_{i,a_i}}{c_{j,M1}}\right)+1_{a_i\in\mathbf{M}}\left(\frac{c_{i,a_i}}{c_{j,M1}}\right)\right)\left(\frac{\widehat{\Pi}_i(\mathbf{A}_n)-\Delta_{i,j}^{M1,P1}(\mathbf{A}_n)}{\widehat{\Pi}_j(\mathbf{A}_n)}\right)\right]\\ &+\sum_{i=j+1}^{N}\left[\left(-1_{a_i\in\mathbf{P}}\left(\frac{r_{i,a_i}}{c_{j,M1}}\right)+1_{a_i\in\mathbf{M}}\left(\frac{c_{i,a_i}}{c_{j,M1}}\right)\right)\left(\frac{\widehat{\Pi}_i(\mathbf{A}_n)+\Delta_{i,j}^{M1,P1}(\mathbf{A}_n)}{\widehat{\Pi}_j(\mathbf{A}_n)}\right)\right]\\ &+\left(\frac{EV(\mathbf{A}_n)}{c_{j,M1}}\right)\left\{\sum_{i=1}^{j-1}\tau_{i,a_i}\left(\frac{\widehat{\Pi}_i(\mathbf{A}_n)-\Delta_{i,j}^{M1,P1}(\mathbf{A}_n)}{\widehat{\Pi}_j(\mathbf{A}_n)}\right)+\tau_{j,P1}+\sum_{i=j+1}^{N}\tau_{i,a_i}\left(\frac{\widehat{\Pi}_i(\mathbf{A}_n)+\Delta_{i,j}^{M1,P1}(\mathbf{A}_n)}{\widehat{\Pi}_j(\mathbf{A}_n)}\right)\right\}.\end{aligned} \tag{11}$$

It should be emphasized that policy $\mathbf{A}_n$ does not have to be a monotone policy, and (11) captures the policy improvement behavior regardless of the type of the policy. The critical ratios can be used to derive conditions under which a control-limit policy is optimal. Reflecting the operating environment of a semiconductor manufacturer, the firm is expected to earn less profit as the process deteriorates, and spend more money on maintenance in deteriorated states. Therefore, let us consider the case that $\frac{r_{j,P1}}{c_{j,M1}}$ is decreasing in j. The relationship between $\frac{r_{j,P1}}{c_{j,M1}}$ with $\gamma_j^{M1,P1}$ in a state j establishes a set of sufficient conditions for an optimal control-limit policy. Specifically, when the decrease in $\gamma_j^{M1,P1}$ as the process condition deteriorates is greater than the decrease in $\frac{r_{j,P1}}{c_{j,M1}}$, the firm is guaranteed to have a monotone optimal policy. This is because the relative values of the critical ratio $\gamma_j^{M1,P1}$ and $\frac{r_{j,P1}}{c_{j,M1}}$ can switch their sign only once. Suppose product $P1$ is manufactured in states 1 through $j-1$, and maintenance action $M1$ is performed in states $j+2$ through N. Let us define the following two policies: $\mathbf{A}_j = [a_1, \ldots, a_{j-1} = P1, a_j, \ldots, a_N = M1]$ and $\mathbf{A}_{j+1} = [a_1, \ldots, a_j = P1, a_{j+1}, \ldots, a_N = M1]$. When the firm considers the switch from maintenance to production in state j, the base policy is $\mathbf{A}_j$, and when it considers the switch in state $j+1$, the base policy is $\mathbf{A}_{j+1}$. The following proposition establishes sufficient conditions for a monotone optimal policy.

Proposition 10. *If the following conditions are satisfied, then there exists a threshold state, $\hat{\jmath}$, such that production is the optimal choice for all states $j < \hat{\jmath}$, and maintenance is optimal for all states $j \geqslant \hat{\jmath}$.*

$$\begin{aligned}\sum_{i=1}^{j}\left[\tau_{i,P1}\Delta_{i,j+1}^{M1,P1}(\mathbf{A}_{j+1})\right] \leqslant &\left[\left[\frac{\sum_{i=1}^{j}[r_{i,P1}\Delta_{i,j+1}^{M1,P1}(\mathbf{A}_{j+1})]+\sum_{i=j+2}^{N}[c_{i,M1}\Delta_{i,j+1}^{M1,P1}(\mathbf{A}_{j+1})]}{EV(\mathbf{A}_{j+1})}\right]\right.\\ &\left.+\sum_{i=j+2}^{N}\left[\tau_{i,M1}\Delta_{i,j+1}^{M1,P1}(\mathbf{A}_{j+1})\right]\right],\end{aligned} \tag{12}$$

$$\begin{aligned}r_{j+1,P1}+c_{j+1,M1} \leqslant &\left[\tau_{j+1,P1}EV(\mathbf{A}_{j+1})-\tau_{j+1,M1}EV(\mathbf{A}_j)\right]\\ &+\left[\left[\sum_{i=1}^{j}\tau_{i,P1}\left(\frac{\widehat{\Pi}_i(\mathbf{A}_{j+1})}{\widehat{\Pi}_{j+1}(\mathbf{A}_{j+1})}\right)+\sum_{i=j+2}^{N}\tau_{i,M1}\left(\frac{\widehat{\Pi}_i(\mathbf{A}_{j+1})}{\widehat{\Pi}_{j+1}(\mathbf{A}_{j+1})}\right)\right]\right.\\ &\left.\left.\times(EV(\mathbf{A}_{j+1})-EV(\mathbf{A}_j))\right]\right],\end{aligned} \tag{13}$$

$$\frac{r_{j,P1}}{c_{j,M1}} \text{ is decreasing in } j. \tag{14}$$

The above proposition proves that if it is optimal to maintain in state $\hat{\jmath}$, then it is optimal to maintain in states $\hat{\jmath}+1$ through N; this fact can greatly reduce the number of potentially optimal policies. In Section 3.1, it has been concluded that the maintenance action is more desirable in deteriorated intermediate states when (1) the value of the change $\Delta_{i,j+1}^{M1,P1}$ in the steady-state expressions is large, (2) the profits and the maintenance costs increase, and (3) when the change in the expected processing times is large. These observations are captured in the sufficient conditions in (12) and (13). Condition (12) states that the total change in the sum of the expected processing times due to the switch from maintenance to production should not be larger the total change that occurs in profits and maintenance costs. Condition (13) focuses on the sum of the profit and the maintenance cost in the state in question, and requires it to be less than or equal to the difference in expected values of the two monotone policies adjusted with normalized expected processing times. Thus, with (12) and (13), any drastic change in profits, maintenance costs, and the expected processing times are prevented, ensuring that the firm does not switch to maintenance and back to production again.

Sufficient conditions in (12)–(14) generalize those reported in the literature. Sloan (2008) provides five sufficient conditions that are required collectively. The conditions can be summarized as: (C1) the profits $r_{j,P1}$ are decreasing in j, and the costs $c_{j,M1}$ are increasing in j; (C2) the machine state has increasing failure rate, i.e., $\sum_{j=l}^{N}p_{ij}^{a}$ is increasing in i for $l = 1, 2, \ldots, N$ and $a \in \{P1, M1\}$; (C3) $c_{j,M1} - r_{j,P1}$ is increasing in j; (C4) for each state l, the sum of the state transition probability matrices is subadditive, i.e., $\sum_{j=l}^{N}p_{ij}^{M1} - \sum_{j=l}^{N}p_{ij}^{P1}$ is decreasing in i for all $l = 1, \ldots, N$; and (C5) the expected completion times are subadditive, i.e., $\tau_{j,M1} - \tau_{j,P1}$ is decreasing in j. Note that (14) is not as restrictive as condition C1 — the profits and the maintenance costs may increase or decrease with respect to j. Conditions (12) and (13) are significantly less restrictive than the subadditivity requirements in their paper, described by conditions C3 and C4. Therefore, the sufficient conditions provided for monotonicity in this paper generalize those reported in the literature.

5.2. Production-related critical ratios

For the production choices, let us consider the policy $\mathbf{A}_n = [a_1, \ldots, a_N]$ with the action $a_j = P1$ in state j. It can be observed that when the firm switches its action from $P1$to $P2$ in state j, the numerator term for state j remains the same, and the numerator terms for all the other states increase; thus, $\delta_i^{P1,P2} > 0$ for all $i = 1, \ldots, N$ where $i \neq j$. As a result of this observation, the critical ratio for production choices in a N-state problem is as follows:

$$\alpha_j^{P1,P2} = \delta_j^{P1,P2} + EV(\mathbf{A}_n)\left(\frac{\tau_{j,P2} - \tau_{j,P1}\delta_j^{P1,P2}}{r_{j,P1}}\right). \tag{15}$$

A non-monotone policy among production choices can be observed when $\frac{r_{i,P2}}{r_{i,P1}} < \alpha_i^{P1,P2}$ in state i and $\frac{r_{j,P2}}{r_{j,P1}} > \alpha_j^{P1,P2}$ in state j where $1 \leqslant i < j \leqslant N$. This implies that the firm prefers to manufacture the low-end product $P1$ with lower profit in a better state i and the high-end product $P2$ with a higher profit in a deteriorated state j. Even if the ratio of profits in each state is constant, an increasing behavior of $\alpha_j^{P1,P2}$ in j can create this scenario. Therefore, it is beneficial to establish the conditions for the increasing and decreasing behavior of the production-related critical ratio in (15).

Proposition 11. *(a) $\alpha_j^{P1,P2}$ is increasing in j when the following three conditions are satisfied: (1) $\delta_j^{P1,P2}$ is increasing in j, (2) $\frac{\tau_{j,P2}}{\tau_{j,P1}} > \delta_j^{P1,P2}$ for each j = 1, …, N, and (3) $\frac{\tau_{j,P2}-\tau_{j,P1}\delta_j^{P1,P2}}{r_{j,P1}}$ is increasing in j; (b) $\alpha_j^{P1,P2}$ is decreasing in j when the following three conditions are satisfied: (1) $\delta_j^{P1,P2}$ is decreasing in j, (2) $\frac{\tau_{j,P2}}{\tau_{j,P1}} < \delta_j^{P1,P2}$ for each j = 1, …, N, and (3) $\frac{\tau_{j,P2}-\tau_{j,P1}\delta_j^{P1,P2}}{r_{j,P1}}$ is decreasing in j.*

The increasing behavior of $\alpha_j^{P1,P2}$ through the above three conditions is useful in establishing a set of sufficient conditions for a monotone optimal policy with respect to production choices. It should be observed that when $r_{j,P1}$ is decreasing in j, $\alpha_j^{P1,P2}$ is increasing in j under conditions 1 and 2, and condition 3 is not necessary as it is automatically satisfied. Considering the operating environment for a semiconductor manufacturer, it can be expected to have the profits decrease as the process deteriorates, and therefore the firm's $\alpha_j^{P1,P2}$ is increasing in j under less restrictive conditions (1 and 2). In the event that $r_{j,P1}$ is increasing in j, $\alpha_j^{P1,P2}$ is still increasing in j under conditions 1, 2, and 3 together. It is obvious that a monotone optimal policy is ensured when $\frac{r_{j,P2}}{r_{j,P1}} < \alpha_j^{P1,P2}$ for all $j = 1, \ldots, N-1$, or when $\frac{r_{j,P2}}{r_{j,P1}} > \alpha_j^{P1,P2}$ for all $j = 1, \ldots, N-1$. Monotonicity is warranted when the ratio of profits is greater than the critical ratio in better states and crosses under the critical ratio only once.

Proposition 12. *The following set of sufficient conditions leads to a monotone policy with respect to production choices: (1) $\frac{r_{j,P2}}{r_{j,P1}}$ is decreasing, (2) $\delta_j^{P1,P2}$ is increasing in j, (3) $\frac{\tau_{j,P2}}{\tau_{j,P1}} > \delta_j^{P1,P2}$ for each j = 1, …, N − 1, and (4) $\frac{\tau_{j,P2}-\tau_{j,P1}\delta_j^{P1,P2}}{r_{j,P1}}$ is increasing in j.*

The above sufficient conditions generalize those reported in the literature significantly. Sloan (2008) reports five sufficient conditions, related to those discussed above (immediately following Proposition 10). In addition to conditions C1 and C2, the following three conditions are required: (C3′) $r_{j,a_j}/\left[1-p_{jj}^{a_j}\right]$ is superadditive for $a_j \in \mathbf{P}$, i.e., $r_{j,P1}/\left[1-p_{jj}^{P1}\right] - r_{j,P2}/\left[1-p_{jj}^{P2}\right]$ is increasing in j; (C4′) for each state l, the sum of the state transition probability matrices is subadditive, i.e., $\sum_{j=l}^{N} p_{ij}^{P1}/\left[1-p_{jj}^{P1}\right] - \sum_{j=l}^{N} p_{ij}^{P2}/\left[1-p_{jj}^{P2}\right]$ is decreasing in i for all $l = 1, \ldots, N$; and (C5′) the expected processing times are subadditive, i.e., $\tau_{j,P1}/\left[1-p_{jj}^{P1}\right] - \tau_{j,P2}/[1-p_{jj}^{P2}]$ is decreasing in j. Proposition 12 does not require anything like condition C1 — the profits may increase or decrease with respect to j. This can be seen in the case when $r_{j,P1}$ and $r_{j,P2}$ are increasing in j with the ratio of profits $\frac{r_{j,P2}}{r_{j,P1}}$ being constant between states; this violates C1. Under the conditions where $\alpha_j^{P1,P2}$ is also constant in each state with a value greater than $\frac{r_{j,P2}}{r_{j,P1}}$, however, our sufficient conditions detect the monotone policy. Moreover, our first condition is less restrictive than condition C3′. Condition C4′ is also more limiting than our third condition. Our second and fourth conditions together are still more general than the subadditivity requirements in their paper. Therefore, the sufficient conditions provided for monotonicity in this paper generalize those reported in the literature. Example 2 provided in the Appendix illustrates a problem for which the sufficient conditions of Sloan (2008) are not met but for which the optimal policy is monotone with respect to the production choices.

5.3. Maintenance-related critical ratios

A similar critical ratio for the maintenance decision can be determined by considering the policy $\mathbf{A}_n = [a_1, \ldots, a_N]$ with the action $a_j = M1$ in state j. It can be observed that when the firm switches its action from $M1$ to $M2$ in state j, the numerator term for state j remains the same, and the numerator terms for all the other states increase, i.e., $\delta_i^{M1,M2} > 0$ for all $i = 1, \ldots, N$ where $i \neq j$. Therefore, the maintenance critical ratio for the N-state problem can be expressed as follows:

$$\lambda_j^{M1,M2} = \delta_j^{M1,M2} + EV(\mathbf{A}_n)\left(\frac{\tau_{j,M1}\delta_j^{M1,M2} - \tau_{j,M2}}{c_{j,M1}}\right). \tag{16}$$

Among maintenance choices, a non-monotone policy can be observed when $\frac{c_{i,M2}}{c_{i,M1}} < \lambda_j^{M1,M2}$ in state i and $\frac{c_{i,M2}}{c_{i,M1}} > \lambda_j^{M1,M2}$ in state j where $1 \leqslant i < j \leqslant N$. This implies that the firm prefers to perform the major maintenance action $M2$ with a higher expense in a better state i and the minor maintenance action $M1$ with a lower cost in a deteriorated state j. Even if the ratio of maintenance costs in each state is constant, an increasing behavior of $\lambda_j^{M1,M2}$ in j can create this scenario. Therefore, it is beneficial to establish the conditions for the increasing/decreasing behavior of the maintenance-related critical ratio.

Proposition 13. *(a)$\lambda_j^{M1,M2}$ is increasing in j when the following three conditions are satisfied: (1) $\delta_j^{M1,M2}$ is increasing in j, (2) $\frac{\tau_{j,M2}}{\tau_{j,M1}} < \delta_j^{M1,M2}$ for each j = 1, …, N, and (3) $\frac{\tau_{j,M1}\delta_j^{M1,M2}-\tau_{j,M2}}{c_{j,M1}}$ is increasing in j; (b) $\lambda_j^{M1,M2}$ is decreasing in j when the following three conditions are satisfied: (1) $\delta_j^{M1,M2}$ is decreasing in j, (2) $\frac{\tau_{j,M2}}{\tau_{j,M1}} > \delta_j^{M1,M2}$for each j = 1, …, N, and (3) $\frac{\tau_{j,M1}\delta_j^{M1,M2}-\tau_{j,M2}}{c_{j,M1}}$ is decreasing in j.*

The decreasing behavior of $\lambda_j^{M1,M2}$ through the above three conditions is useful in establishing a set of sufficient conditions for a monotone policy with respect to maintenance choices. It should be observed that when $c_{j,M1}$ is increasing in j, $\lambda_j^{M1,M2}$ is decreasing in j under conditions 1 and 2 (of part b), and condition 3 is not necessary as it is automatically satisfied. Considering the operating environment for a semiconductor manufacturer, maintenance costs can be expected to increase as the process deteriorates, and therefore the firm's $\lambda_j^{M1,M2}$ is decreasing in j under less restrictive conditions (1 and 2). In the event that $c_{j,M1}$ is decreasing in j, $\lambda_j^{M1,M2}$ is still decreasing in j under conditions 1, 2, and 3 together. It is easy to observe that a monotone policy is ensured when $\frac{c_{j,M2}}{c_{j,M1}} < \lambda_j^{M1,M2}$ (or when $\frac{c_{j,M2}}{c_{j,M1}} > \lambda_j^{M1,M2}$) for all $j = 2, \ldots, N$. Once again, monotonicity is warranted when $\frac{c_{j,M2}}{c_{j,M1}}$ is greater than $\lambda_j^{M1,M2}$ in better states and crosses under the critical ratio only once.

Proposition 14. *The following set of sufficient conditions leads to a monotone policy with respect to maintenance choices: (1) $\frac{c_{j,M2}}{c_{j,M1}}$ is decreasing in j, (2) $\delta_j^{M1,M2}$ is increasing in j, (3) $\frac{\tau_{j,M2}}{\tau_{j,M1}} < \delta_j^{M1,M2}$ for each j = 1, …, N, and (4) $\frac{\tau_{j,M1}\delta_j^{M1,M2} - \tau_{j,M2}}{c_{j,M1}}$ is increasing in j.*

The above sufficient conditions generalize those reported in the literature significantly. Sloan (2008) extends the maintenance policy results of Hopp and Wu (1990), and requires collectively: (C1) the costs $c_{j,M1}$ and $c_{j,M2}$ are non-decreasing in j, (C2) the machine state has increasing failure rate, (C3) $c_{i,a_i \in \mathbf{M}}$ is superadditive, i.e., the difference in the costs $c_{j,M1} - c_{j,M2}$ is increasing in j, (C4) for state l, the sum of the state transition probability matrices is subadditive, i.e., $\sum_{j=l}^{N}\left[p_{ij}^{M2} - p_{ij}^{M1}\right]$ is decreasing in i for all $l = 1, \ldots, N$, and (C5) the expected maintenance times are subadditive, i.e., $\tau_{j,M1} - \tau_{j,M2}$ is decreasing in j. The conditions listed in Proposition 14 are much more general. Condition C3 is similar to our first condition; however, ours is less restrictive. Similarly, condition C4 is similar — but more restrictive — than our second and fourth conditions combined. In the Appendix, Example 3 illustrates the situation in which some of the Sloan (2008) conditions are not met but for which the optimal policy is monotone with respect to the maintenance actions.

Let us define $\hat{\jmath}_{P1,M1}$ as the minor maintenance threshold with respect to the standard product $P1$ and $\hat{\jmath}_{P1,M2}$ as the major maintenance threshold with respect to $P1$. Proposition 14 is equivalent to saying that $\hat{\jmath}_{P1,M1} \leqslant \hat{\jmath}_{P1,M2}$; this fact can reduce the set of potentially optimal policies drastically. Specifically, the decision maker does not have to consider all four actions in each of the $N - 2$ states. In states $\hat{\jmath}_{P1,M2}$ through N, for example, the choice is restricted to be between $M1$ and $M2$, and in states between $\hat{\jmath}_{P1,M1}$ and $\hat{\jmath}_{P1,M2}$ the choice is restricted to $P1$, $P2$ or $M1$.

The above results highlight the value of the critical ratios. Determining the optimal solution for a given problem is fairly straightforward. For example, a standard linear programming formulation for an SMDP can be used. The real leverage from the critical ratios, especially in the presence of production requirements, is the ability to narrow the set of potentially optimal policies. Once this reduced set of policies is identified, then the firm's short-term production and maintenance decisions are greatly simplified. Although we do not specify a solution algorithm here, the critical ratios can be used to develop heuristics to streamline scheduling decisions.

6. Conclusions

This paper considers a manufacturer's production and maintenance choices under deteriorating process conditions. The firm has to make three decisions in each machine state: (1) whether to produce or maintain the process, (2) if production is chosen, which product to manufacture, and (3) if maintenance is elected, whether to employ a major or a minor maintenance action. Each of these three decisions has trade-offs. In the first, production speeds the process deterioration, and maintenance is likely to improve it; however, while production earns profits, maintenance leads to a cost. As deterioration takes place, if the firm chooses production over maintenance, it commits to future maintenance costs with higher probability. The choice of the product also influences the process deterioration: a high-end product provides a higher profit, but takes longer to manufacture and accelerates the process deterioration; thus, it elevates the need for future maintenance and its associated costs. A low-end product brings less profit, but has a lower probability of process deterioration, and leads to smaller probabilities of maintenance needs and associated costs. The maintenance choice influences the process improvement similarly. In a deteriorated equipment state, if the firm chooses to maintain the system, then it incurs a direct cost of the maintenance action and an indirect cost associated with the loss of profits that can be gained from manufacturing its products. A major maintenance action incurs a higher direct cost, but is more likely to improve the process than a minor maintenance which costs less money. As a result, a major maintenance can result in a higher yield of products manufactured in less deteriorated states, resulting in higher overall profits. Excessive maintenance, however, can actually reduce net throughput by devoting more time to maintenance rather than production. The paper develops a model that captures the complex relationships between these three decisions, process deterioration and improvement probabilities, profits and costs, and the expected processing times. We incorporate market demand considerations by including minimum and maximum production requirements for each product and examine how these requirements influence the optimal policy.

The paper makes four sets of contributions. First, it develops three critical ratios. The first critical ratio determines whether the firm should manufacture or maintain the equipment. The second critical ratio enables the firm to choose the preferred product in each state. The third critical ratio informs the decision maker about the appropriate maintenance action. These critical ratios have economic interpretations. The first two critical ratios can be interpreted as *reservation prices*, i.e., the maximum amount of money the decision maker should be willing to pay in order to switch from maintenance to production in the first, and from a low-end product to a high-end product in the second. The third critical ratio enables the decision maker to establish an upper bound on the cost of the major maintenance action corresponding to the *maximum* amount of money she should be willing to spend. Second, the paper shows how these three critical ratios can be combined in order to determine the optimal action among all possible choices. The combination enables the firm to capture the above trade-offs *simultaneously*. These critical ratios are then generalized to problem settings that feature an arbitrary number of machine states. Third, the paper demonstrates the influence of production requirements on the choice of the optimal policy and the critical ratios used to determine the optimal solution. It shows that, depending on the length of its expected time, maintenance can play a strategic role in increasing the throughput of a high demand product. Fourth, a set of sufficient conditions are developed that lead to monotone optimal policies. These conditions are demonstrated to be significantly more generalized than those reported in the literature. Monotone policies suggest that the firm manufacture its high-technology products in better process conditions and switch to low-technology products as the machine deteriorates. At some level of deterioration, production is no longer viable, and maintenance is performed. The firm should employ a minor maintenance with continued deterioration, and perform a major maintenance in significantly deteriorated states.

The initial motivation for the model was to explore short-term decision making regarding production and maintenance decisions. Once the critical ratios are computed for a given scenario, determining the optimal policy is straightforward. In Section 4, we discussed how demand-related throughput requirements actually narrow the possibilities and thus make the search for an optimal policy faster. In addition to these operational-level decisions, the model can also provide insights relevant to longer-term, strategic decisions. For example, interpreting the critical ratios as reservation prices can inform decisions about product mix (e.g., is it more profitable to narrow the product scope), pricing, and product substitution (a common practice in the semiconductor industry). One might also gain insight into process technology choices — e.g., can the current generation of equipment be used profitably as the product technology advances (meaning greater circuit density and therefore higher sensitivity to equipment condition)? Although not specifically formulated for these types of decisions, examination of such strategic issues is an interesting area for future research.

Acknowledgments

This research is partially funded by the Robert H. Brethen Operations Management Institute at Syracuse University.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ejor.2012.11.052.

References

Aghezzaf, E.H., Jamali, M.A., Ait-Kadi, D., 2007. An integrated production and preventive maintenance planning model. European Journal of Operational Research 181, 679–685.
Batun, S., Maillart, L.M., 2012. Reassessing tradeoffs inherent to simultaneous maintenance and production planning. Production and Operations Management 21, 396–403.
Ben-Daya, M., 2002. The economic production lot-sizing problem with imperfect production processes and imperfect maintenance. International Journal of Production Economics 76, 257–264.
Boone, T., Ganeshan, R., Guo, Y.M., Ord, J.K., 2000. The impact of imperfect processes on production run times. Decision Sciences 31, 773–787.
Cassady, C.R., Kutanoglu, E., 2005. Integrating preventive maintenance planning and production scheduling for a single machine. IEEE Transactions on Reliability 54, 304–309.
Chakraborty, T., Giri, B., Chaudhuri, K., 2008. Production lot sizing with process deterioration and machine breakdown. European Journal of Operational Research 185, 606–618.
Dehayem Nodem, F., Kenne, J., Gharbi, A., 2009. Hierarchical decision making in production and repair/replacement planning with imperfect repairs under uncertainties. European Journal of Operational Research 198, 173–189.
El-Ferik, S., 2008. Economic production lot-sizing for an unreliable machine under imperfect age-based maintenance policy. European Journal of Operational Research 186, 150–163.
Gilbert, S.M., Bar, H.M., 1999. The value of observing the condition of a deteriorating machine. Naval Research Logistics 46, 790–808.
Gilbert, S.M., Emmons, H., 1995. Managing a deteriorating process in a batch production environment. IIE Transactions 27, 233–243.
Groenevelt, H., Pintelon, L., Seidmann, A., 1992a. Production batching with machine breakdowns and safety stocks. Operations Research 40, 959–971.
Groenevelt, H., Pintelon, L., Seidmann, A., 1992b. Production lot sizing with machine breakdowns. Management Science 38, 104–123.
Hadidi, L.A., Al-Turki, U.M., Rahim, A., 2012. Integrated models in production planning and scheduling, maintenance and quality: a review. International Journal of Industrial and Systems Engineering 10, 21–50.
Hariga, M., Ben-Daya, M., 1998. The economic manufacturing lot-sizing problem with imperfect manufacturing processes: bounds and optimal solutions. Naval Research Logistics 45, 423–433.
Hopp, W.J., Wu, S.-C., 1988. Multiaction maintenance under Markovian deterioration and incomplete state information. Naval Research Logistics 35, 447–462.
Hopp, W.J., Wu, S.-C., 1990. Machine maintenance with multiple maintenance actions. IIE Transactions 22, 226–232.
Iravani, S.M.R., Duenyas, I., 2002. Integrated maintenance and production control of a deteriorating production system. IIE Transactions 34, 423–435.
Karamatsoukis, C., Kyriakidis, E., 2010. Optimal maintenance of two stochastically deteriorating machines with an intermediate buffer. European Journal of Operational Research 207, 297–308.
Kazaz, B., Sloan, T., 2008. Production policies for multi-product systems with deteriorating process condition. IIE Transactions 40, 187–205.
Kim, C.H., Hong, Y., Chang, S.Y., 2001. Optimal production run length and inspection schedules in a deteriorating production process. IIE Transactions 33, 421–426.
Lee, C.Y., 2004. Machine scheduling with an availability constraint. In: Leung, J.Y.-T. (Ed.), Handbook of Scheduling: Algorithms, Models, and Performance Analysis. CRC Press, Boca Raton, FL.
Lee, H.L., Rosenblatt, M.J., 1987. Simultaneous determination of production cycle and inspection schedules in a production system. Management Science 33, 1125–1136.
Lee, H.L., Rosenblatt, M.J., 1989. A production and maintenance planning model with restoration cost dependent on detection delay. IIE Transactions 21, 368–375.
Lee, J.S., Park, K.S., 1991. Joint determination of production cycle and inspection intervals in a deteriorating production system. Journal of the Operational Research Society 42, 775–783.
Liao, G., Chen, Y.H., Sheu, S., 2009. Optimal economic production quantity policy for imperfect process with imperfect repair and maintenance. European Journal of Operational Research 195, 348–357.
Makis, V., 1998. Optimal lot sizing and inspection policy for an EMQ model with imperfect inspections. Naval Research Logistics 45, 165–186.
Makis, V., Fung, J., 1998. An EMQ model with inspections and random machine failures. Journal of the Operational Research Society 49, 66–76.
Nourelfath, M., 2011. Service level robustness in stochastic production planning under random machine breakdowns. European Journal of Operational Research 212, 81–88.
Porteus, E.L., 1986. Optimal lot sizing, process quality improvement and setup cost reduction. Operations Research 34, 137–144.
Porteus, E.L., 1990. The impact of inspection delay on process and inspection lot sizing. Management Science 36, 999–1007.
Puterman, M.L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York.
Rosenblatt, M.J., Lee, H.L., 1986. Economic production cycles with imperfect production processes. IIE Transactions 18, 48–54.
Sana, S.S., 2010a. An economic production lot size model in an imperfect production system. European Journal of Operational Research 201, 158–170.
Sana, S.S., 2010b. A production-inventory model in an imperfect production process. European Journal of Operational Research 200, 451–464.
Sloan, T.W., 2004. A periodic review production and maintenance model with random demand, deteriorating equipment, and binomial yield. Journal of the Operational Research Society 55, 647–656.
Sloan, T.W., 2008. Simultaneous determination of production and maintenance schedules using inline equipment condition and yield information. Naval Research Logistics 55, 116–129.
Sloan, T.W., Shanthikumar, J.G., 2000. Combined production and maintenance scheduling for a multiple-product, single-machine production system. Production and Operations Management 9, 379–399.
Sloan, T.W., Shanthikumar, J.G., 2002. Using in-line equipment and yield information for maintenance scheduling and dispatching in semiconductor wafer fabs. IIE Transactions 34, 191–209.
Su, C.-T., Wu, S.-C., Chang, C.-C., 2000. Multiaction maintenance subject to action-dependent risk and stochastic failure. European Journal of Operational Research 125, 133–148.
Wang, H., 2002. A survey of maintenance policies of deteriorating systems. European Journal of Operational Research 139, 469–489.
Yano, C.A., Lee, H.L., 1995. Lot sizing with random yields: a review. Operations Research 43, 311–334.