

# Production policies under deteriorating process conditions

BURAK KAZAZ<sup>1,\*</sup> and THOMAS W. SLOAN<sup>2</sup>

<sup>1</sup>Whitman School of Management, Syracuse University, Syracuse, NY 13204, USA  
E-mail: bkazaz@syr.edu

<sup>2</sup>College of Management, University of Massachusetts Lowell, Lowell, MA 01854, USA  
E-mail: Thomas\_Sloan@uml.edu

Received February 2006 and accepted January 2007

---

This paper examines a single-stage production system that manufactures multiple products under deteriorating equipment conditions. The machine condition worsens with production, and improves with maintenance. The condition of the process can be in any one of several discrete states, and transitions from state to state follow a semi-Markov process. In many production environments, the quality or yield of output depends heavily on the condition of the production process. The problem considers the trade-offs between manufacturing products that have a higher profit, a longer processing time, and therefore, a higher deterioration probability versus products that have a smaller profit, shorter processing time with a lower process deterioration probability. The firm needs to determine the optimal production choice in each state in a way that maximizes the long-run expected average reward per unit time.

The paper makes three sets of contributions. First, it introduces the concept of critical ratios for the firm's manufacturing decision at each state regarding whether to switch from one product to another. Second, through the use of critical ratios, the main result shows that the optimal production choice for each state can be determined *independently* of the actions taken in other states, despite the complex interconnections between the production decisions and state transitions. Third, the paper provides generalizations that illustrate the depth, scope and richness of the proposed solution technique by extending the model in the number of machine states, to settings where maintenance is performed in intermediate states, and to settings where transition probabilities are influenced by both mean and variance of processing times.

**Keywords:** Production policies, process deterioration, semi-Markov decision process

## 1. Introduction

This paper studies optimal production decisions for a manufacturing system that produces multiple products under deteriorating equipment conditions. In many production environments, the quality or output yield depends heavily on the condition of the production process. Traditionally, researchers exploring the connections between process condition and yield have focused on quantity, i.e., how large should be the production batches given that some fraction of the finished units will be defective. In contrast, this paper considers the question of *which* product should be produced depending on the process condition.

We consider a single-stage manufacturing system that produces multiple product types. The condition of the system deteriorates with production, and the quality (yield) of the final output is a function of both the process condition and the product type. In environments as diverse as semiconductor wafer manufacturing, pharmaceutical manufacturing and optical lens production, the process condi-

tion deteriorates over time and reduces the amount of yield. The goal of the production manager is to determine the optimal production decision, i.e., *which* product to produce, in each equipment condition. When the process reaches the worst state, the manufacturer performs maintenance and returns the equipment to its best state.

The process condition deteriorates differently according to which product is being produced. In the motivating example, we consider a semiconductor wafer manufacturer who is concerned with the production of two products: a high-end and a low-end technology product. A high-end product typically has more circuitry per unit area on the computer chip than a low-end product and therefore requires a longer processing time. As production takes place, the equipment becomes more contaminated, resulting in a higher level of process deterioration. It is expected that a high-end product will earn more revenue than a low-end product. However, the higher circuit density means that a high-end product will have a lower yield than a low-end product for a given process condition (i.e., level of contamination). Furthermore, the longer processing time of a high-end product increases the likelihood of process deterioration. Therefore, the manager's trade-off at each

---

\*Corresponding author

equipment condition is: produce the high-end product and earn a higher revenue but increase the risk of process deterioration versus produce the low-end product and earn a lower revenue but reduce the risk of process deterioration. Although the high-end product brings in more revenue per unit, it also increases the likelihood of more frequent maintenance, and thus will increase the overall maintenance cost.

The main focus of this research is to identify the structural properties of the optimal production policy. The study characterizes all potentially optimal solutions and determines the conditions that make them optimal.

The paper makes three sets of contributions. First, it introduces the concept of the *critical ratio* of revenues, under which the decision-maker is indifferent in her/his choice of the product to be manufactured. Thus, it is sufficient for the firm to compare the revenues of products in each state with the critical ratio of the state in order to determine the optimal production policy. The critical ratios have significant managerial implications because they enable the manufacturer to evaluate analytically the *reservation price* of a product, i.e., the minimum that she/he needs to earn in order to justify the profitable production of this product over other products. Owing to the fact that machine deterioration probabilities change with the production choices made in all states, it is unexpected to have *separable* optimal production decisions for each state. A rather surprising and counter-intuitive result, the second set of contributions shows that despite the interdependencies between processing times, machine deterioration probabilities and production choices, the optimal production decision in a state can be made *independently* of the production choices made in other states. The third set of contributions involves generalizing the problem to more complex settings and illustrates the depth and richness of the proposed solution technique. When the problem is extended in the number of states describing equipment condition, for example, the number of potentially optimal policies increases dramatically. In our approach, however, the decision-maker needs to evaluate only one additional set of critical ratios for this new state in order to determine the optimal production policy. In another extension, when maintenance is allowed in intermediate states, we show the condition that when it is optimal to perform maintenance in a state, then it is optimal to do so in all of the following (worse) states. It is then proven that the general problem with many states can be reduced to a problem setting that considers the maintenance action only in the threshold state and production in prior (better) states. The final extension illustrates how these critical ratios can be evaluated in more complex settings: first when both the mean and variance of the processing times influence machine state transition probabilities, and then in the absence of a functional relationship between processing times and machine deterioration probabilities.

The results of the paper have both operational and managerial implications. Operationally, they facilitate the development of intuitive and easy-to-implement policies. Man-

agerially, they shed light on decisions regarding product mix, pricing and process technology.

## 2. Literature review

A great deal of research has been performed on production systems with variable yield. Readers are referred to the extensive survey by Yano and Lee (1995) for a complete review of the various issues and approaches used to study such problems. The research that is most relevant to our problem is the subset of variable yield models that explicitly accounts for the interaction between process condition and yield. The first models in this area are Porteus (1986) and Rosenblatt and Lee (1986). In both of these papers, the classical Economic Manufacturing Quantity (EMQ) model is extended to account for changes in the process condition. Specifically, the process begins in an “in-control” state with perfect production quality, and after some time, shifts to an “out-of-control” state in which some fraction of production is defective. The process state is observable only at the end of a production run. Both papers show that the optimal production quantity is smaller than the quantity resulting from the traditional EMQ approach.

Many variations of these early models have been pursued. Some models examine different cost structures that depend on when defective items are detected (Lee and Rosenblatt, 1989; Lee and Park, 1991). Other models allow inspections (i.e., observation of the process state) during production, and the decision about when to inspect is optimized along with the production quantity (Lee and Rosenblatt, 1987; Porteus, 1990; Kim *et al.*, 2001). The problem has also been extended to incorporate various aspects of maintenance and reliability such as preventive maintenance (Zequeira *et al.*, 2004), machine failures (Makis and Fung, 1998; Boone *et al.*, 2000) and imperfect maintenance (Ben-Daya, 2002).

All of the aforementioned models only consider single-product systems or treat all products the same way. While this may be appropriate in some contexts, in many environments different products will be affected differently by the equipment condition. For example, leading-edge technology products are more sensitive to the process condition. The case of multiple products where the yield depends on the equipment condition was first examined by Sloan and Shanthikumar (2000, 2002). However, both papers assume that the processing times are equal for products, resulting in equal machine deterioration probabilities. Thus, the machine deterioration probabilities do not depend on the choice of the product. Sloan and Shanthikumar (2002) applies the results of the earlier paper in a heuristic fashion to a multi-stage environment. Products make multiple visits to each workstation (referred to as “layers”), so the total manufacturing time of different products may be different. However, the processing times at each station are assumed to be the same, so even though

one product may require 20 visits to each station (i.e., 20 layers) and another product requires only ten visits to each station (i.e., ten layers), the model only accounts for this difference with respect to the expected rewards, and not with respect to the processing times or transition probabilities. As we shall see in the forthcoming model section, these differences between products play an important role in determining the optimal production policy. Although they derive sufficient conditions on the rewards that ensure monotone production policies (i.e., policies that call for the production of high-end products in better states and low-end products in worse states), they fail to provide any structural results regarding the optimality conditions using differing processing times and transition probabilities.

Our paper departs from earlier studies in two ways. First, products are differentiated not only based on their yield (and reward) but also based on their processing times and their impact on the equipment deterioration process. Therefore, for a given state, the transition probabilities vary according to the product choice. Second, while the majority of previous research has been focused on the question of *how much* to produce, this paper investigates the question of *which* product to produce. In addition, those papers that do investigate which products to produce only consider sufficient conditions for optimality of certain types of policies, such as monotone policies. Our work is a significant generalization of these papers as it develops the *necessary* and sufficient conditions to characterize *all* forms of optimal policies (monotone and non-monotone), while capturing the complex interdependencies between processing times, deterioration probabilities and rewards.

### 3. The model

This section presents the model used to prescribe the firm's production decisions in a single-stage manufacturing system. The firm can produce multiple products, indexed by parameter  $k = 1, 2, \dots, K$ , corresponding to a total of  $K$  products. The equipment condition deteriorates as production takes place. Each product influences the process deterioration differently, therefore, the firm's objective is to determine the set of optimal production decisions that maximize the long-run expected average reward. The analysis of this section isolates the impact of varying expected processing times on the machine deterioration under equal variances (in processing times); the case of unequal variances is examined in Section 4.

The equipment condition is described by a set of  $N$  discrete states, and is indexed by  $i$  (and  $j = 1, 2, \dots, N$ ), where  $i = 1$  represents the best state and  $i = N$  represents the worst state. At each decision epoch, the firm is forced to make a two-part decision: first, whether to produce or maintain; and second, if production is picked, *which product* to produce. When the firm chooses to produce, the action is denoted by variable  $a \in \{1, 2, \dots, K\}$ , and when the firm decides to maintain the equipment (so that the process returns

to its best state) is represented by  $a = m$ . The time required to perform action  $a$  is a random variable with mean  $\tau_a$  and variance  $\sigma_a^2$ .

The transition probability for the process is denoted as  $p_{ij}^a$ , corresponding to the probability of the equipment being in state  $j$  at the next decision epoch given that at the current epoch the machine is in state  $i$  and action  $a$  is taken. It should be observed here that the transition probabilities are defined in such a way that the machine condition generally gets worse while producing, but would not move to a better state. More precisely, the transition probabilities for production actions ( $a = 1, 2, \dots, K$ ) are defined as follows:

$$p_{ij}^a \begin{cases} > 0 & \text{for all } 1 \leq i \leq j \leq N, \\ = 0 & \text{for all } j < i, \\ = 1 & \text{for } j = i = N. \end{cases}$$

For the maintenance action (corresponding to action  $a = m$ ), the equipment returns to the best state with probability one:

$$p_{ij}^m \begin{cases} = 0 & \text{for all } 1 \leq i \leq N \text{ and } 2 \leq j \leq N, \\ = 1 & \text{for all } i = 1, \dots, N \text{ and } j = 1. \end{cases}$$

It should be noted here that even though the transition probabilities defined as  $p_{ij}^a$  refer to the machine state only at decision points, the equipment condition can change between decision epochs. For example, even if production is commenced in state 1, the machine condition may deteriorate significantly during production, before the action is completed.

The process deterioration probabilities are impacted differently by the choice of production action  $a_i = 1, \dots, K$  in each state  $i$ . It is the motivating argument of this paper that the longer the expected production time for a product, the higher the deterioration probability for that action. Therefore, the relationship between the process deterioration probabilities of two different products, products  $k$  and  $l$ , is defined as a function of the relative values of the expected (mean) processing times and the variance in processing times (denoted by  $\sigma_k^2$  and  $\sigma_l^2$  for products  $k$  and  $l$ , respectively):

$$p_{ij}^k = c\beta_{k,l}p_{ij}^l + \varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2) \text{ for all } 1 \leq i < j \leq N, \quad (1)$$

where  $\beta_{k,l} = \tau_k/\tau_l$  is the ratio of expected production times for products  $k$  and  $l$ ,  $c$  is a constant that indicates how the deterioration probabilities change *vis-à-vis* the ratio of expected processing times and  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)$  is the functional term representing the impact of the variances on processing times. When  $c > 1$ , the deterioration probability for product  $k$  increases at a rate faster than the ratio of expected processing times, and when  $1/\beta_{k,l} < c < 1$ , it increases at a rate slower than the ratio of expected production times. The difference between the variances of the two products also influences the transition probabilities, and this is expressed by the function  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)$ . The value of  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)$  can be positive or negative, corresponding to an increment or reduction in the transition probability, and is restricted to

be such that  $|\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)| < \min(c\beta_{k,l}p_{ij}^l, 1 - c\beta_{k,l}p_{ij}^l)$ . The term  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)$  can be interpreted as the *variance effect* in the change in deterioration probabilities. In this section, emphasis is placed on the impact of the expected processing times, so it is assumed that  $\sigma_k^2 = \sigma_l^2$  for products  $k$  and  $l$ , and therefore  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2) = 0$ . The impact of the variance effect is studied in depth in Section 4.

The choice of the product to be manufactured not only influences the process deterioration probabilities, but also the reward earned in each state. This is because each product brings a different reward in each state. As the machine deteriorates with production actions, the yield for each product decreases, leading to reduced rewards. Therefore, the reward for each product  $k$  is non-increasing in the machine state, i.e.,  $r_{1k} \geq r_{2k} \geq \dots \geq r_{Nk}$ .

This study examines the interrelationships and interdependencies of the three problem parameters, namely the *production times* which impact the *machine deterioration probabilities* and the *rewards* earned in each state with production. In order to capture the trade-offs between expected processing times and rewards, the products are rank ordered according to their expected processing times:  $\tau_1 \geq \tau_2 \geq \dots \geq \tau_K$ . Thus, product 1 has the longest expected processing time, and product  $K$  has the shortest expected processing time. In order to study the relationship between the rewards and the processing times, the products are assumed to have their rewards in the following order in the best state:  $r_{11} \geq r_{12} \geq \dots \geq r_{1K}$ , where product 1 (which has the longest expected processing time) provides the highest reward and product  $K$  (which has the lowest expected processing time) has the lowest reward in state 1. It should be noted here that there are no assumptions made about the ordering of rewards in other states. We next define  $RR_{k,l}^i = r_{ik}/r_{il}$  as the ratio of rewards between products  $k$  and  $l$  in a given state  $i$ . Finally, it should be stated here that we make the mild assumption that the rewards are such that the long-run average reward for each policy featuring the production of a single product type has a positive value; otherwise this product type is not profitable and its production would not be justified.

To summarize, the time between decision epochs, the machine state transition probabilities, and the rewards earned depend only on the current state and the action taken. Thus, this scenario can be modeled as a Semi-Markov Decision Process (SMDP). A stationary policy (i.e., time invariant) induces a discrete-time Markov chain that characterizes the equipment condition at decision epochs. This is referred to as the *Embedded Markov Chain* (EMC). The transition probabilities defined above describe the evolution of the EMC over time; that is,  $p_{ij}^a = \Pr\{X_{t+1} = j \mid X_t = i, a_t = a\}$ , where  $X_t$  denotes the machine state and  $a_t$  denotes the action taken at decision epoch  $t$ .

Several approaches are available to solve this type of problem (see Howard (1960), Heyman and Sobel (1984) or Puterman (1994) for general discussions and Tijms (1986)

for SMDP-specific material). We use a *policy improvement* approach, which finds the optimal decision rule by starting with a reference policy and comparing it to another policy that differs by only one action in one state. To accomplish this, one must compute the expected long-run average profit of a given policy, and we denote this expected value as  $EV$ .

Let  $\mathbf{A} = [a_i \mid i = 1, \dots, N]$  denote a stationary policy vector that specifies action  $a_i$  when the machine is in state  $i$ . Define  $\pi_i(\mathbf{A})$  as the stationary (or *steady-state*) probability that the associated EMC is in state  $i$  when policy  $\mathbf{A}$  is used. A unique set of steady-state probabilities is guaranteed as long as the EMC induced by a stationary policy results in a single, closed set of recurrent states. The conditions shown in (Tijms, 1986) are satisfied because the number of machine states is finite, production causes the equipment condition to deteriorate and maintenance causes the machine to return to the best state. Thus, there exists a single set of recurrent states, and therefore there exists a unique set of steady-state probabilities, regardless of the initial state of the process. Note that the stationary probability for one state may depend on the machine state transition probabilities of *all* other states, so  $\pi_i(\mathbf{A})$  is a function of the entire policy vector. However, the rewards and the production (and maintenance) times depend only on the action taken in the current state; thus, they do not depend on the entire policy vector. The average reward rate of policy  $\mathbf{A}$  can then be expressed as

$$EV(\mathbf{A}) = \frac{\sum_{i=1}^N r_{i,a_i} \pi_i(\mathbf{A})}{\sum_{i=1}^N \tau_{a_i} \pi_i(\mathbf{A})}. \quad (2)$$

A policy  $\mathbf{A}^*$  is average reward optimal if  $EV(\mathbf{A}^*) \geq EV(\mathbf{A})$  for each stationary policy  $\mathbf{A}$ . The optimal action in state  $i$  is defined as  $a_i^*$ .

The total number of policies one can generate in this problem grows significantly in the number of products and in the number of states that describe the process condition. Considering  $K$  products and  $N$  machine states with maintenance being performed only in the worst state, the manufacturer has to evaluate the expected values of  $(K)^{N-1}$  potentially optimal policies before choosing the one that maximizes the expected average reward. The purpose of this paper is to explore the structural properties of the problem by using the approach outlined above to provide insight into the solution. For this purpose, we exploit the analytical properties of a smaller setting of the original problem with two products and three machine states. The detailed analysis of this smaller setting forms the foundation of the generalizations that follow. All proofs are provided in the Appendix.

### 3.1. The core problem

The core of the analysis regarding the optimal production choice can be developed using two products in a setting which defines the machine condition in three states. In

this simplified version of the problem, we consider policies that require production actions in the first two states (describing a better machine condition with higher yields and revenues) and the maintenance action in state 3. Using the earlier notation, action  $a_i = 1$  refers to producing product 1, action  $a_i = 2$  refers to producing product 2 in state  $i = 1, 2$  and action  $a_3 = m$  corresponds to performing maintenance in state 3. Four policies are possible in this context:  $\mathbf{A}_1 = [1, 1, m]$ : produce product 1 in states 1 and 2;  $\mathbf{A}_2 = [1, 2, m]$ : produce product 1 in state 1, and product 2 in state 2;  $\mathbf{A}_3 = [2, 1, m]$ : produce product 2 in state 1, and product 1 in state 2; and  $\mathbf{A}_4 = [2, 2, m]$ : produce product 2 in states 1 and 2. Given that production is to be undertaken in states 1 and 2, we can now focus on the question of which product to produce in these two states. Although the problem can be solved computationally, the goal here is to characterize the optimal policy without explicitly solving the problem each time.

We begin our analysis by describing the steady-state probabilities for machine states. Steady-state probabilities for a policy  $\mathbf{A}$  can be determined by using the machine state transition probabilities associated with the action taken in each state. Due to the fact that processing times for different products are different, the machine state transition probabilities depend on which product is produced in each state. Making use of the state balance equations for the EMC, the stationary probability for states 1, 2 and 3 associated with policy  $\mathbf{A} = [a_1, a_2, a_3 = m]$ , which specifies that action  $a_1$  is taken in state 1 and  $a_2$  in state 2, are

$$\pi_1(\mathbf{A}) = \frac{(1 - p_{22}^{a_2})}{(1 - p_{22}^{a_2}) + p_{12}^{a_1} + (1 - p_{11}^{a_1})(1 - p_{22}^{a_2})},$$

$$\pi_2(\mathbf{A}) = \frac{p_{12}^{a_1}}{(1 - p_{22}^{a_2}) + p_{12}^{a_1} + (1 - p_{11}^{a_1})(1 - p_{22}^{a_2})},$$

and

$$\pi_3(\mathbf{A}) = \frac{(1 - p_{11}^{a_1})(1 - p_{22}^{a_2})}{(1 - p_{22}^{a_2}) + p_{12}^{a_1} + (1 - p_{11}^{a_1})(1 - p_{22}^{a_2})}.$$

Note that a change in one action in one state changes *all* of the stationary probabilities, therefore making it difficult to compare different production policies. Thus, one would not expect the optimal production choice in a state to be independent of the decisions made in other states. This motivates the investigation of optimality conditions that account for the best action to be taken in each state.

The expected value of a particular policy can be determined by plugging the above stationary probabilities into Equation (2) and simplifying:

$$\begin{aligned} EV(\mathbf{A} = [a_1, a_2, a_3 = m]) &= \frac{r_{1,a_1}(1 - p_{22}^{a_2}) + r_{2,a_2}p_{12}^{a_1} + r_{3m}(1 - p_{11}^{a_1})(1 - p_{22}^{a_2})}{\tau_{a_1}(1 - p_{22}^{a_2}) + \tau_{a_2}p_{12}^{a_1} + \tau_m(1 - p_{11}^{a_1})(1 - p_{22}^{a_2})}. \end{aligned} \quad (3)$$

The following example identifies the ‘‘common sense’’ approaches that are widely used in developing solution approaches for similar problems. It demonstrates, however, that these approaches do not necessarily generate optimal policies.

*Example 1.* Consider the following problem with two products and three machine states. The profit earned for product 1 is  $r_{11} = 950$  in state 1 and  $r_{21} = 600$  in state 2, and the profit for product 2 is  $r_{12} = 600$  in state 1 and  $r_{22} = 301$  in state 2. The maintenance cost is  $r_{3m} = -800$ . The expected time required to produce product 1 is  $\tau_1 = 2$ , and product 2 is  $\tau_2 = 1$ , yielding a ratio of processing times  $\beta_{1,2} = \tau_1/\tau_2 = 2$ . The expected time required to perform maintenance is  $\tau_m = 2$ . We let  $c = 0.95$ , which means that the deterioration probabilities for product 1 increase at a lower rate than the ratio of expected processing times. The deterioration probabilities for product 1 are then equal to  $p_{ij}^1 = c\beta_{1,2}p_{ij}^2 = 1.90 \times p_{ij}^2$  for all  $1 \leq i < j \leq 3$ . Performing maintenance, on the other hand, returns the equipment condition to state 1 with probability one. The machine state transition probability for each action,  $p_{ij}^a$ , refers to the probability of the machine being in state  $j$  at the next decision epoch given that at the current epoch the machine is in state  $i$  and action  $a$  is taken. Their values are

$$[p_{ij}^1] = \begin{bmatrix} 0.430 & 0.285 & 0.285 \\ 0 & 0.430 & 0.570 \\ 0 & 0 & 1 \end{bmatrix},$$

$$[p_{ij}^2] = \begin{bmatrix} 0.700 & 0.150 & 0.150 \\ 0 & 0.700 & 0.300 \\ 0 & 0 & 1 \end{bmatrix},$$

$$[p_{ij}^m] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

To determine the optimal policy, one could easily substitute the appropriate values into Equation (3) for each policy, and choose the one with the highest expected value. But is there a way to determine the optimal policy without explicitly comparing all policies? One might intuit that, for example, choosing the action that maximizes the expected reward for each state would be optimal. Based on these rewards, producing product 1 in states 1 and 2 appears to be optimal using this ‘‘greedy’’ approach, so the optimal policy would be  $[1, 1, m]$ , generating an expected value of 191.787 for its average reward. A second common sense approach to determining the optimal policy might be to choose the action that maximizes the expected average reward *per unit time* for each state. In state 1, producing product 2 earns  $600/1 = 600$  per unit time, and product 1 earns  $950/2 = 475$ . In state 2, producing product 2

earns  $301/1 = 301$ , while product 1 earns only  $600/2 = 300$ . Thus, the second approach indicates that policy  $[2, 2, m]$  would be optimal, generating an expected value of 195.476 for its average reward.

It turns out that both of these “common sense” approaches are incorrect: the optimal policy is  $[2, 1, m]$  with an expected value of 196.535 for its average reward. This result is quite surprising. One would think that if product 2 is superior to product 1 in state 1, then it would also be superior in state 2. Similarly, if product 1 is preferred to product 2 in a worse state, then it should also be preferred to product 2 in a better state. We now turn our attention to explaining this counter-intuitive behavior by exploring the structural properties of the optimal policy.

We next introduce a solution approach that uses the comparison of the policies that feature the manufacturing of a single product,  $\mathbf{A}_1 = [1, 1, m]$  and  $\mathbf{A}_4 = [2, 2, m]$ , corresponding to the production of only product 1 and only product 2, respectively. The solution approach takes one of these two products as its *reference product* and the policy that features the manufacturing of this product in each state as the *reference policy*. We begin our analysis by considering product 2 as the reference product and policy  $\mathbf{A}_4 = [2, 2, m]$  as the reference policy. We next investigate the conditions that make the firm switch its production choice from this reference product (product 2) to product 1 in each state. Let us first examine state 2, the second-to-last state. Consider policy  $\mathbf{A}_3 = [2, 1, m]$ , which differs from  $\mathbf{A}_4$  only in that product 1 is produced in state 2 rather than product 2. Referring to Equation (3), this means that  $a_2 = 2$  for  $\mathbf{A}_4$ , while  $a_2 = 1$  for  $\mathbf{A}_3$ . Comparing these two policies determines when the firm prefers to switch its manufacturing choice from product 2 (the reference product) to product 1 in state 2.

Is it possible to find the point at which the decision-maker is indifferent between the two products in state 2? Such a point depends on the ratio of the rewards earned for each product in state 2, and thus we refer to the indifference point as the *critical ratio* for the firm’s decision to switch its manufacturing choice from the reference product (product 2) to product 1. We define  $\underline{\alpha}_{k,l}^i$  as the *critical ratio* of the rewards in state  $i$  for products  $k$  and  $l$  when product  $l$  is the reference product. When the actual ratio of rewards in state  $i$  for these two products, defined earlier by  $RR_{1,2}^i$ , is greater than  $\underline{\alpha}_{1,2}^i$ , then the firm prefers to produce product 1 rather than product 2. Otherwise, if  $RR_{1,2}^i$  is less than the critical ratio  $\underline{\alpha}_{1,2}^i$ , then the firm prefers to keep manufacturing product 2 rather than switching to product 1. Therefore, the comparison of policies  $\mathbf{A}_4 = [2, 2, m]$  and  $\mathbf{A}_3 = [2, 1, m]$  in the core problem leads to the critical ratio of rewards in state 2, and is expressed as  $\underline{\alpha}_{1,2}^2$ . Similarly, the firm can develop the critical ratio for state 1,  $\underline{\alpha}_{1,2}^1$ , by comparing the reference policy  $\mathbf{A}_4 = [2, 2, m]$  with  $\mathbf{A}_2 = [1, 2, m]$ , where the two policies differ only in the production decision made in state 1.

A different set of *critical ratios* can be obtained by comparing the new reference policy of  $\mathbf{A}_1 = [1, 1, m]$  with policy  $\mathbf{A}_2 = [1, 2, m]$  (where the two policies differ only in state 2) and with  $\mathbf{A}_3 = [2, 1, m]$  (where the two policies differ only in state 1). These critical ratios correspond to the ratio of rewards that the firm prefers to switch from the reference product of product 1 to product 2. They are denoted by  $\bar{\alpha}_{k,l}^i$  for products  $k$  and  $l$  in state  $i$  when product  $k$  is the reference product. When the actual ratio of rewards in state  $i$  for these two products (denoted by  $RR_{1,2}^i$ ) is greater than  $\bar{\alpha}_{1,2}^i$ , then the firm prefers to keep manufacturing product 1 rather than switching to product 2. Otherwise, when  $RR_{1,2}^i$  is less than the critical ratio  $\bar{\alpha}_{1,2}^i$ , then the firm prefers to switch its manufacturing from product 1 to product 2.

It should be observed here that one of the two single-product policies will be preferred. When  $EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m])$ , the firm can use  $\underline{\alpha}_{1,2}^i$  values as its active set of critical ratios. Otherwise, when  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ , then the firm uses  $\bar{\alpha}_{1,2}^i$  in order to choose the product to be manufactured in state  $i$ . We define  $\alpha_{1,2}^i$  as the *active critical ratio*, and it is determined by the relative value of the two single-product production policies; i.e.,  $\alpha_{1,2}^i = \bar{\alpha}_{1,2}^i$  when  $EV(\mathbf{A}_1) > EV(\mathbf{A}_4)$ , and  $\alpha_{1,2}^i = \underline{\alpha}_{1,2}^i$  when  $EV(\mathbf{A}_1) \leq EV(\mathbf{A}_4)$ . The following proposition provides the closed-form expressions for the critical ratios in each state, which correspond to the exact ratio of rewards to determine which product is preferred for manufacturing in each state.

**Proposition 1.** *There exists a set of critical ratios for each state that determines the firm’s manufacturing preference in each state:*

$$\bar{\alpha}_{1,2}^i = c\beta_{1,2} + \beta_{1,2}(1 - c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{i2}} \quad \text{for each state } i = 1, 2, \quad (4)$$

$$\underline{\alpha}_{1,2}^i = c\beta_{1,2} + \beta_{1,2}(1 - c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{i2}} \quad \text{for each state } i = 1, 2, \quad (5)$$

and

$$\alpha_{1,2}^i = \begin{cases} \bar{\alpha}_{1,2}^i & \text{when } EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m]) \\ \underline{\alpha}_{1,2}^i & \text{when } EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m]) \end{cases} \quad \text{for each state } i = 1, 2. \quad (6)$$

(i) When  $RR_{1,2}^i > \alpha_{1,2}^i$ , the firm prefers to manufacture product 1 in state  $i$ ; (ii) when  $RR_{1,2}^i < \alpha_{1,2}^i$ , the firm prefers to manufacture product 2 in state  $i$ ; and (iii) when  $RR_{1,2}^i = \alpha_{1,2}^i$ , the firm is indifferent between manufacturing products 1 and 2 in state  $i$ .

It should be observed that the critical ratios defined by  $\bar{\alpha}_{1,2}^i$  and  $\underline{\alpha}_{1,2}^i$  have similar expressions, however, their values are different unless the expected values of the two

single-product policies are equal, i.e., when  $EV(\mathbf{A}_1) = EV(\mathbf{A}_4)$ . As evident from Equations (4) and (5), the critical ratios are impacted by the same set of parameters: (i) the rate that the deterioration probabilities are influenced by the increase in processing times; (ii) the expected processing times and their relative ratios; (iii) the rewards earned in each state; and (iv) the expected value of the reference policy.

The critical ratios have the same behavior in response to the changes in parameters. First, it can be observed that the value of each critical ratio differs from one state to another. For example, the critical ratio  $\alpha_{1,2}^1$  for state 1 is not equal to the critical ratio  $\alpha_{1,2}^2$  for state 2 unless the rewards for the reference product are identical in both states, i.e.,  $r_{1a} = r_{2a}$ . The equality of rewards corresponds to the situation when the machine deterioration does not decrease the yield in both states. However, the premise of this problem is that the yields, and therefore the rewards, decrease as the machine condition deteriorates. Therefore, the case of equal rewards is not of interest in this context. Secondly, the increasing (or decreasing) behavior of the critical ratios depends on the value of  $c$ , the rate that the deterioration probabilities increase with respect to expected processing times. These observations are formalized in the following two propositions:

**Proposition 2.** (i) When  $c \geq 1$ , the critical ratio  $\alpha_{1,2}^i$  is non-increasing in  $i$ ; and (ii) when  $1/\beta_{12} \leq c < 1$ , the critical ratio  $\alpha_{1,2}^i$  is non-decreasing in  $i$ .

**Proposition 3.** (i) When  $c \geq 1$  and  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ , then  $\underline{\alpha}_{1,2}^i > \bar{\alpha}_{1,2}^i$ ; (ii) when  $c \geq 1$  and  $EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m])$ , then  $\underline{\alpha}_{1,2}^i \leq \bar{\alpha}_{1,2}^i$ ; (iii) when  $1/\beta_{12} \leq c < 1$  and  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ , then  $\underline{\alpha}_{1,2}^i < \bar{\alpha}_{1,2}^i$ ; and (iv) when  $1/\beta_{12} \leq c < 1$  and  $EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m])$ , then  $\underline{\alpha}_{1,2}^i \geq \bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ .

The critical ratio defined by  $\alpha_{1,2}^i$  provides managerial insight using economic principles. Considering the reference policy of producing the low-end technology product (i.e., product 2), for example, the critical ratio  $\alpha_{1,2}^i$  prescribes the *reservation price* for product 1; that is, the critical ratio multiplied by the reward of product 2 is the *minimum* amount of money that a manager should earn in order to justify the production of a higher-end technology (i.e., product 1). Thus, when the actual ratio of rewards is larger than the critical ratio, the firm benefits more by manufacturing product 1. However, when the actual ratio of rewards is less than the critical ratio  $\alpha_{1,2}^i$ , the firm benefits more by manufacturing product 2.

We next introduce a unique solution approach to solving the production planning problem for multiple products under deteriorating process conditions. The following theorem prescribes the optimal policy with the use of the critical

ratio characterizing the optimal production decision in each state.

**Theorem 1.** *The optimal production decision in each state can be determined independently by comparing the actual ratio of rewards ( $RR_{1,2}^i$ ) with the critical ratio of  $\alpha_{1,2}^i$ .*

(i) When  $RR_{1,2}^i \geq \alpha_{1,2}^i$ , then  $a_i^* = 1$ ; and (ii) when  $RR_{1,2}^i < \alpha_{1,2}^i$ , then  $a_i^* = 2$ ; and (iii) when  $\alpha_{1,2}^i = \bar{\alpha}_{1,2}^i$ , it is never the case that  $RR_{1,2}^i < \bar{\alpha}_{1,2}^i$  for both states  $i = 1, 2$  at the same time; and when  $\alpha_{1,2}^i = \underline{\alpha}_{1,2}^i$ , it is never the case that  $RR_{1,2}^i < \underline{\alpha}_{1,2}^i$  for both states  $i = 1, 2$  at the same time.

The consequence of the above theorem is that the optimal production policy can be determined easily once the expected average rewards for the two reference policies are computed. More importantly, despite the interdependencies between the steady-state probabilities in Equation (2), the optimal production choice for each state is *independent* of the choices made in other states. Put differently, the manufacturing choices in each state are *separable* despite the interdependencies between the processing times, the deterioration probabilities and the rewards earned with production decisions.

*Continuation of Example 1:* The expected values of the single-product policies are  $EV(\mathbf{A}_1 = [1, 1, m]) = 191.787 < EV(\mathbf{A}_4 = [2, 2, m]) = 195.476$ . Therefore, the ratios of rewards are compared with the critical ratios of  $\alpha_{1,2}^i = \underline{\alpha}_{1,2}^i$  in each state  $i = 1, 2$ . Since  $RR_{1,2}^1 = 950/500 = 1.900 < \alpha_{1,2}^1 = 1.939$ , product 2 is the optimal choice in state 1. In state 2,  $RR_{1,2}^2 = 600/301 = 1.993 > \alpha_{1,2}^2 = 1.965$ , so product 1 is the optimal choice. This confirms that the optimal policy is  $[2, 1, m]$ , as stated above. Using the values of  $\alpha_{1,2}^i$  for  $i = 1, 2$ , one can calculate how much profit would be required for product 1 to make it the optimal choice in a particular state. This corresponds to the *reservation price* of the manufacturer in order to choose product 1 over product 2. Specifically, since  $\alpha_{1,2}^1 = 1.939$ , one would require a profit of 969.5 ( $= 500 \times 1.939$ ) to prefer product 1 over product 2 in state 1.

Theorem 1 provides further managerial insight into the manufacturer's production choices. It is the motivating application of this paper that when the machine condition deteriorates, the production yield decreases, resulting in lower rewards in worse states. Consider the case when the yields of both products decrease at the same rate for a given increase in equipment deterioration. Thus, the ratio of rewards would be constant between states for the two products, i.e.,  $RR_{1,2}^i$  is constant for all states. In this scenario of equal yield (and reward) reduction, the following proposition proves that the firm switches its manufacturing choice at most once between products. The switch depends on the value of  $c$ , the rate that deterioration probabilities are influenced by the ratio of expected processing times.

**Proposition 4.** *When  $RR_{1,2}^i$  is constant for each  $i$ , then the firm switches its optimal production choice at most once: (i) if*

$c \geq 1$  and  $a_1^* = 2$ , then if  $a_i^* = 1$  for some  $i < N$ , then  $a_j^* = 1$  for all  $j > i$ ; (ii) If  $1/\beta_{12} < c < 1$  and  $a_1^* = 1$ , then if  $a_i^* = 2$  for some  $i < N$ , then  $a_j^* = 2$  for all  $j > i$ .

The above proposition provides insight into the monotonicity behavior of the optimal policy. An increasing monotone policy is such that the production choice starts with the manufacturing of the high-end technology product (e.g., product 1) and switches to low-end products (e.g., product 2) as the machine condition worsens, but does not switch back to high-end products. For example, a policy such as  $\mathbf{A} = [1, 1, 2, m]$  is an increasing monotone policy since it features the manufacturing of product 1 in better states (states 1 and 2), and switches to product 2 in state 3, but does not switch back to product 1 again. The above proposition proves that when the deterioration probabilities increase at a rate slower than the increase in processing times, i.e.,  $1/\beta_{12} < c < 1$ , the firm's optimal policy is strictly an increasing monotone policy under the case of equal yield and reward reductions. On the other hand, a decreasing monotone policy is such that the production choice starts with the manufacturing of the low-end technology product (e.g., product 2) and switches to high-end products (e.g., product 1) as the machine condition worsens, but does not switch back to low-end products. For example, a policy such as  $\mathbf{A} = [2, 1, 1, m]$  is a decreasing monotone policy since it features the manufacturing of product 2 in the best state (state 1), and switches to product 1 in states 2 and 3, but does not switch back to product 2 again. The above proposition proves that when the deterioration probabilities increase at a rate faster than the increase in processing times, i.e.,  $c \geq 1$ , the firm's optimal policy is strictly a decreasing monotone policy under the case of equal yield and reward reductions. Thus, a non-monotone policy such as  $\mathbf{A} = [2, 1, 2, m]$  cannot be optimal in the case of equal yield and reward reductions.

### 3.2. The solution technique for the setting with N machine states

The analysis presented in the previous section shows that the production decisions in each state can be made independently of the actions taken in other states. This separable decision-making technique is originally proven in a setting that features only three machine states, but can be easily extended to a setting with  $N$  machine states. In the new setting with  $N$  machine states, the critical ratios, defined by  $\bar{\alpha}_{k,l}^i$  and  $\underline{\alpha}_{k,l}^i$ , continue to be useful in determining the optimal production decision in each state while providing managerial insight. They can be expressed as

$$\bar{\alpha}_{1,2}^i = c\beta_{1,2} + \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A} = [1, \dots, 1, m])}{r_{i2}} \text{ for each state } i = 1, \dots, N,$$

$$\underline{\alpha}_{1,2}^i = c\beta_{1,2} + \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A} = [2, \dots, 2, m])}{r_{i2}} \text{ for each state } i = 1, \dots, N,$$

and

$$\alpha_{1,2}^i = \begin{cases} \bar{\alpha}_{1,2}^i & \text{when } EV(\mathbf{A} = [1, \dots, 1, m]) > EV(\mathbf{A} = [2, \dots, 2, m]) \\ \underline{\alpha}_{1,2}^i & \text{when } EV(\mathbf{A} = [1, \dots, 1, m]) \leq EV(\mathbf{A} = [2, \dots, 2, m]) \end{cases} \text{ for each state } i = 1, \dots, N.$$

Using the same approach utilized to prove Theorem 1 (specifically, the proof by induction and by contradiction), the optimal production decision in each state can be determined independently by comparing the actual ratio of rewards ( $RR_{1,2}^i$ ) with the active critical ratio  $\alpha_{1,2}^i$ .

**Corollary 1.** (i) When  $RR_{1,2}^i \geq \alpha_{1,2}^i$ , then  $a_i^* = 1$ ; and (ii) when  $RR_{1,2}^i < \alpha_{1,2}^i$ , then  $a_i^* = 2$ .

It should also be observed that when  $EV(\mathbf{A} = [1, \dots, 1, m]) > EV(\mathbf{A} = [2, \dots, 2, m])$ , the ratio of rewards in each state cannot be smaller than the corresponding critical ratio in all states, and  $RR_{1,2}^i < \bar{\alpha}_{1,2}^i$  does not hold true for all states  $i = 1, \dots, N$  at the same time. Similarly, when  $EV(\mathbf{A} = [1, \dots, 1, m]) \leq EV(\mathbf{A} = [2, \dots, 2, m])$ , the ratio of rewards in each state cannot be higher than the corresponding critical ratio in all states, i.e.,  $RR_{1,2}^i > \underline{\alpha}_{1,2}^i$  does not hold true for all states  $i = 1, \dots, N$  at the same time. Once again, the optimal production decision in each state can be made independently of the decisions made in other states.

The solution approach presented for the two-product problem can be extended to problem settings with three or more products. For example, when there are three products, i.e.,  $k = 1, 2$  and  $3$ , the solution technique features two, at most three, pairwise comparisons in order to determine the optimal production decision in each state.

### 3.3. The impact of maintenance in intermediate states

Our primary interest in studying this problem is to analyze production policies. Nevertheless, one may wish to consider the possibilities of other maintenance policies. For example, in the core problem we compared policies  $[2, 2, m]$  and  $[2, 1, m]$ . But what about a policy such as  $[2, m, m]$ ? Can the optimality of a policy that calls for maintenance only in the worst state be guaranteed unless other maintenance policies are considered?

When maintenance is allowed to be performed in an intermediate state in a problem setting with  $K$  products and  $N$  machine states, there are a total of  $\sum_{i=1}^{N-1} (K)^i$  policies that induce an EMC with a single, closed set of recurrent states. While this increases the complexity of the problem significantly, as detailed below, the manufacturer does not need to enumerate them all before choosing the one that maximizes the expected average reward. We first establish that the maintenance policy is a *control-limit* policy. In other words, there exists a threshold state,  $\hat{i}$ , such that if maintenance is optimal in state  $\hat{i}$ , then maintenance is also optimal



in all states  $i \geq \hat{i}$ . The following lemma, based on a result from Kao (1973), specifies sufficient conditions for a control-limit maintenance policy.

**Lemma 1.** *When for each  $l = 1, \dots, N$ ,  $\sum_{i=1}^l p_{ij}^a$  is non-increasing in  $i$  for all actions  $a$ , there exists a threshold state  $\hat{i}$  such that if maintenance is optimal in state  $\hat{i}$ , then it is optimal in all states  $i$ , where  $\hat{i} \leq i \leq N$ .*

It should be observed here that the maintenance cost is considered to be equal between states in deriving the above lemma. However, one might argue that the maintenance cost might be increasing in the state number, corresponding to higher maintenance costs for worse states. Although not proven here, the above lemma can be extended easily to a problem to accommodate for an increasing maintenance cost in proving the existence of a threshold state.

The threshold state  $\hat{i}$  is useful in re-establishing the critical ratios developed for the production actions. For the machine states  $i < \hat{i}$ , the optimal production decisions can still be determined by the use of the critical ratios. However, these ratios require the information in which state  $\hat{i}$  of  $N$  machine states, the maintenance actions are taken because the reference policy needs to be adjusted for the maintenance actions in states  $\hat{i}$  through  $N$ . Therefore, we revise the notation for these critical ratio expressions in order to accommodate maintenance actions between states  $\hat{i}$  through  $N$ . Let us now define  $\bar{\alpha}_{k,l}^i(N, M)$  and  $\underline{\alpha}_{k,l}^i(N, M)$  as the *critical ratios* of the rewards in state  $i$  for products  $k$  and  $l$  associated with an  $N$ -state problem setting when using a policy that calls for production in states  $1, \dots, M - 1$  and maintenance in states  $M, M + 1, \dots, N$ :

$$\begin{aligned} \bar{\alpha}_{k,l}^i(N, M) &= c\beta_{k,l} + \beta_{k,l}(1 - c) \\ &\times \frac{\tau_l EV(\mathbf{A} = [a_1 = k, \dots, a_{M-1} = k, a_M = m, \dots, a_N = m])}{r_{il}}, \\ \underline{\alpha}_{k,l}^i(N, M) &= c\beta_{k,l} + \beta_{k,l}(1 - c) \\ &\times \frac{\tau_l EV(\mathbf{A} = [a_1 = l, \dots, a_{M-1} = l, a_M = m, \dots, a_N = m])}{r_{il}}. \end{aligned}$$

The solution techniques prescribed in this paper continue to hold even under this revision, because the following proposition greatly reduces the effort needed to obtain the optimal solution. It shows that the active critical ratio expression can be simplified by considering information regarding the threshold state, i.e., the first state where maintenance is performed.

**Proposition 5.**  $\alpha_{k,l}^i(N, M) = \alpha_{k,l}^i(N - 1, M) = \dots = \alpha_{k,l}^i(M, M)$  for all states  $i = 1, \dots, M - 1$ .

The significance of the above proposition is that the analysis of a problem with maintenance in intermediate states can be reduced to a setting with a smaller number of states. By using the induction approach, the proposition shows that the critical ratios that accommodate maintenance actions in states  $M$  through  $N$  are identical to the critical

ratios obtained for the problem setting with  $M$  machine states and maintenance being performed only in the last state. Alternatively, the expected value of a policy for an  $N$ -state problem that has production actions in states 1 through  $M - 1$ , and maintenance in states  $M$  through  $N$ , is equal to the expected value of a policy for a  $M$ -state problem that has production actions in states 1 through  $M - 1$ , and maintenance in state  $M$ . The consequence of this result is that the problem that has maintenance actions in intermediate states can be reduced to the problem setting with maintenance being performed only in the last state. This result, once again, validates the solution technique proposed earlier for the production planning problem for multiple products under deteriorating process conditions.

#### 4. The behavior of critical ratios

In Section 3, we investigated the effect of differing mean processing times on the optimal product choice. The critical ratios were developed under the assumption that  $\sigma_k^2 = \sigma_l^2$  and therefore the variance term  $\varepsilon_{ij}^{kl}(\sigma_k^2, \sigma_l^2)$  in Equation (1) was equal to zero. In this section, we assume that  $\sigma_k^2 \neq \sigma_l^2$ , allowing us to investigate the impact of processing time variance on these critical ratios, and thus, on the optimal product choice.

##### 4.1. The impact of processing times variance

Let us consider the core problem of Section 3.1 with three machine states, but this time with three products. Product 3 has the shortest expected processing time with the lowest variance of processing times, and earns the smallest reward in the best machine state. When compared with product 3, product 2 has a higher expected processing time and equal variance, and it earns a higher reward in the best state. The mean processing time of product 1 is equal to that of product 2 (no mean effect between products 1 and 2), but it has a different variance than product 2. To summarize, we have  $r_{11} > r_{12} > r_{13}$ ,  $\tau_1 = \tau_2 > \tau_3$  and  $\sigma_1^2 \neq \sigma_2^2 = \sigma_3^2$ . While the comparison of products 2 and 3 highlights the impact of the mean of the processing times (as in Section 3.1), the comparison of products 1 and 2 isolates the impact of variance on the critical ratios. The comparison of products 1 and 3 incorporates both the mean and variance effects on these critical ratios.

The variance in processing times influences the transition probabilities. Using Equation (1), the transition probabilities between products 1 and 2 can be expressed as  $p_{ij}^1 = p_{ij}^2 + \varepsilon_{ij}^{12}(\sigma_1^2, \sigma_2^2 \mid \sigma_1^2 \neq \sigma_2^2)$ , where the sum of the variance terms for each initial state is equal to zero, i.e.,  $\sum_{j \geq i} \varepsilon_{ij}^{12}(\sigma_1^2, \sigma_2^2) = 0$  for  $i = 1, 2$ . Using the same approach detailed in Section 3, a comparison of products 1 and 2 leads

to the following critical ratios that highlight the impact of the variance in processing times:

$$\alpha_{1,2}^1 = 1 - \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{12}(1-p_{22}^a)}(r_{2a_2} - \tau_2 EV[a_1, a_2, m]) + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{12}}(r_{3m} - \tau_m EV[a_1, a_2, m]), \quad (7)$$

$$\alpha_{1,2}^2 = 1 + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{22}p_{12}^a}(r_{1a_1} - \tau_2 EV[a_1, a_2, m]) + (1-p_{11}^a)(r_{3m} - \tau_m EV[a_1, a_2, m]), \quad (8)$$

where  $EV[a_1, a_2, m]$  is the expected value of the reference policy with actions  $a_1$  in state 1 and  $a_2$  in state 2. As before,  $\alpha_{1,2}^i$  represents the active critical ratio; that is, when  $EV[1, 1, m] > EV[2, 2, m]$ , for example, we have  $\alpha_{1,2}^i = \bar{\alpha}_{1,2}^i$ , and  $a_1 = a_2 = 1$ .

Understanding the sign of the variance terms in Equation (1) sheds more light on the behavior of the critical ratios in Equations (7) and (8). For convenience, we consider the case with increasing variance in processing times, and study the above example when product 1 has a higher variance than product 2, i.e.,  $\sigma_1^2 > \sigma_2^2$  (and  $\tau_1 = \tau_2$ ). It is generally expected that the probability of remaining in a state when product 1 is manufactured is less than when product 2 is produced; thus, let us assume  $p_{ii}^1 < p_{ii}^2$  for  $i = 1, 2$ . In this case, the variance term  $\varepsilon_{ii}^{12}(\sigma_1^2, \sigma_2^2 | \sigma_1^2 > \sigma_2^2)$  becomes negative for each state  $i = 1, 2$ . Similarly, increasing variance generally implies that the probability of reaching the worst state is expected to be higher when product 1 is manufactured than when product 2 is produced; therefore, we assume that  $p_{iN}^1 \geq p_{iN}^2$  for  $i = 1, 2$ . This means  $\varepsilon_{i3}^{12}(\sigma_1^2, \sigma_2^2 | \sigma_1^2 > \sigma_2^2) \geq 0$  for  $i = 1, 2$ . Note that the variance term for the deterioration probability  $\varepsilon_{ii}^{12}(\sigma_1^2, \sigma_2^2 | \sigma_1^2 > \sigma_2^2) < -\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)$  can still be positive or negative. Under these assumptions, we can now provide more insight into the increasing (or decreasing) behavior of Equations (7) and (8). Note that when Equations (7) and (8) are greater than one, the firm requires a higher reward in order to justify the production of the product with a different processing time variance. It is already known that  $r_{3m} - \tau_m EV[a_1, a_2, m] < 0$ , and  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(r_{3m} - \tau_m EV[a_1, a_2, m]) > 0$ . Consider the case when the variance of product 1 is slightly higher than that of product 2, such that  $\varepsilon_{13}^{12}(\sigma_1^2, \sigma_2^2) = 0$  and  $\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2) = -\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2) > 0$ . In this case, both critical ratios are strictly increasing in each state when  $r_{ia_i} - \tau_2 EV[a_1, a_2, m] < 0$ ; thus, the firm needs a higher reward in each state to justify the manufacture of the product with a higher variance. On the other hand, when  $r_{ia_i} - \tau_2 EV[a_1, a_2, m] > 0$  for each action in each state, the increasing (or decreasing) behavior of the critical ratios depends on the relative values of  $(r_{ia_i} - \tau_2 EV[a_1, a_2, m])$  and  $(1-p_{ii}^a)(r_{3m} - \tau_m EV[a_1, a_2, m]) < 0$ . The behavior is determined by the reward that can be earned in the deteriorated state relative to the further deterioration probability times the maintenance cost, where

the latter can be interpreted as a simplified expected maintenance expense. A similar observation can be made when the variance of product 1 is significantly higher, such that  $\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2) = 0$  and  $\varepsilon_{13}^{12}(\sigma_1^2, \sigma_2^2) = -\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2) > 0$ . In this case, the critical ratio for state 1 is strictly increasing because its value is greater than one. The critical ratio for state 2 is strictly increasing when  $r_{1a_1} - \tau_2 EV[a_1, a_2, m] < 0$ ; and is decreasing if  $r_{1a_1} - \tau_2 EV[a_1, a_2, m] > (1-p_{11}^a)|r_{3m} - \tau_m EV[a_1, a_2, m]|$ . In sum, managers typically need to earn a higher reward in the best state in order to justify the manufacture of a product with higher processing time variance. In deteriorated machine states, however, the necessary reward to justify the manufacture of a product with high variance depends on the relative value of the reward earned and the (expected) maintenance expense. The following proposition summarizes the necessary and sufficient conditions for the behavior of these critical ratios when  $p_{ii}^1 \leq p_{ii}^2$  and  $p_{i3}^1 \geq p_{i3}^2$  for  $i = 1, 2$ .

**Proposition 6.** *Increasing variance in processing times implies: (i)  $\alpha_{1,2}^1$  is increasing if  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1-p_{22}^a)(r_{3m} - \tau_m EV[a_1, a_2, m]) > \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)(r_{2a_2} - \tau_2 EV[a_1, a_2, m])$ , otherwise it is decreasing; (ii)  $\alpha_{1,2}^2$  is increasing if  $r_{1a_1} - \tau_2 EV[a_1, a_2, m] < -(1-p_{11}^a)(r_{3m} - \tau_m EV[a_1, a_2, m])$ , otherwise it is decreasing.*

The combined effect of mean and variance on the critical ratios integrates the above results with those presented in Section 3.1. The critical ratios obtained through the comparison of products 1 and 3 are

$$\bar{\alpha}_{1,3}^1 = \bar{\alpha}_{2,3}^1 - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1-p_{11}^a)}[c\beta_{2,3}(r_{13} - \tau_2 EV[2, 2, m]) + (1-p_{11}^a)(r_{3m} - \tau_m EV[2, 2, m])], \quad (9)$$

$$\bar{\alpha}_{1,3}^2 = \bar{\alpha}_{2,3}^2 + \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^a}c\beta_{2,3}(r_{23} - \tau_2 EV[2, 2, m]) - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1-p_{11}^a)}{r_{23}p_{12}^3}(r_{3m} - \tau_m EV[2, 2, m]), \quad (10)$$

$$\underline{\alpha}_{1,3}^1 = \underline{\alpha}_{2,3}^1 - \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1-p_{22}^a)}(r_{23} - \tau_3 EV[3, 3, m]) + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}}(r_{3m} - \tau_m EV[3, 3, m]), \quad (11)$$

$$\underline{\alpha}_{1,3}^2 = \underline{\alpha}_{2,3}^2 + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^3}[(r_{13} - \tau_3 EV[3, 3, m]) + (1-p_{11}^a)(r_{3m} - \tau_m EV[3, 3, m])]. \quad (12)$$

Similar observations can be made regarding the behavior of the critical ratios. The combined effect of the mean and variance can be easily seen in the values of  $\bar{\alpha}_{1,3}^i$  when product 2 is the reference product. It should be observed that the variance terms in Equations (11) and (12) developed for  $\underline{\alpha}_{1,3}^i$  follow the same behavioral pattern as the terms in

Equations (7) and (8), except that they consider the rewards and transition probabilities of product 3 rather than product 2. The following proposition provides the necessary and sufficient conditions for the behavior of the critical ratios relative to the mean and variance of processing times when  $p_{ii}^1 \leq p_{ii}^2$  and  $p_{i3}^1 \geq p_{i3}^2$  for  $i = 1, 2$ .

**Proposition 7.** (i) In state 1, the firm needs to earn a higher reward to switch from product 3 to product 1 than from product 3 to product 2 when  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{22}^3)(r_{3m} - \tau_m EV[3, 3, m]) > \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)(r_{23} - \tau_3 EV[3, 3, m])$ ; (ii) in state 2, the firm needs to earn a higher reward to switch from product 3 to product 1 than from product 3 to product 2 when  $r_{13} - \tau_3 EV[3, 3, m] < -(1 - p_{11}^3)(r_{3m} - \tau_m EV[3, 3, m])$ ; (iii) in state 1, the firm needs to earn a higher reward to switch from product 1 to product 3 than from product 2 to 3 when  $c\beta_{2,3}(r_{13} - \tau_2 EV[2, 2, m]) > -(1 - p_{22}^2)(r_{3m} - \tau_m EV[2, 2, m])$ ; (iv) in state 2, the firm needs to earn a higher reward to switch from product 1 to product 3 than from product 2 to 3 when  $\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)c\beta_{2,3}(r_{23} - \tau_2 EV[2, 2, m]) > \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(r_{3m} - \tau_m EV[2, 2, m])$ .

#### 4.2. The most general form of critical ratios

Until now, the relationship between the processing times and the deterioration probabilities has been defined as in Equation (1). In this section, we develop the critical ratio expressions in the absence of a specific relationship as in Equation (1), corresponding to the analysis under the presence of arbitrary state transition probabilities. In this case, the deterioration probability for a product with a longer expected processing time can be smaller than that of a product with a shorter expected processing time. Thus, the new expressions developed here correspond to the most general form of the critical ratios. To facilitate the expression and explanation of the critical ratios for the setting with  $N$  machine states, we define the following three parameters:  $\theta_{k,l}(i) = (1 - p_{ii}^k)/(1 - p_{ii}^l)$  is the ratio of exit probabilities from state  $i$  for products  $k$  and  $l$ ; this can also be perceived as the ratio of the sum of deterioration probabilities for products  $k$  and  $l$  when the machine is in state  $i$ , for all products  $1 \leq k < l \leq K$  and  $N - 1$ ;  $\delta_{k,l}(i, i + j) = p_{i,i+j}^k/(1 - p_{i,i+j}^l)$  is the ratio of the  $j$ -step deterioration (transition) probability of product  $k$  when the machine is in state  $i$  to the sum of the deterioration probabilities for product  $l$  when the machine goes to state  $i + j$ , for all products  $1 \leq k < l \leq K$  and states  $1 \leq j \leq N - 1 - i$  and  $1 \leq i \leq N - 1$ ; and,  $\eta_k(i, i + j) = p_{i,i+j}^k/(1 - p_{i,i+j}^k)$  is the ratio of the  $j$ -step deterioration probability when the machine is in state  $i$  for product  $k$  with respect to the sum of deterioration probabilities when the machine goes to state  $i + j$ , for all  $1 \leq k \leq K$  and  $1 \leq j \leq N - 1 - i$  and  $1 \leq i \leq N - 1$ . Using these new parameters one can develop the most general form of the critical ratio expressions for the case when the deterioration probability of a product is

not defined as a function of the processing time. The critical ratios that help the firm determine whether to produce product  $k$  or  $l$  in state  $i = N - j$  where  $j \leq N - 1$  are expressed as follows:

$$\begin{aligned} \bar{\alpha}_{k,l}^{(N-j)} &= \theta_{k,l}(N-j) + (\beta_{k,l} - \theta_{k,l}(N-j)) \frac{\tau_l EV(\mathbf{A} = [k, \dots, k, m])}{r_{(N-j),l}} \\ &+ \sum_{i=1}^{j-1} \left( \sum_{s=1}^i \left( \begin{array}{c} (\theta_{k,l}(N-j)\delta_{l,k}(N-j, N-j+s) \\ - \eta_k(N-j, N-j+s)) \\ \times \left( 1_{s=i} + \sum_{u=1}^{j-1-s} \left( \prod_{x=N-j+s}^{N-1-u} \eta_k(x, x+u) \right) \right) \right) \right) \right) \\ &\times \left( \frac{r_{N-j+i,l} - \tau_l EV(\mathbf{A} = [k, \dots, k, m])}{r_{N-j,l}} \right) \end{array} \right) \quad (13) \\ \underline{\alpha}_{k,l}^{(N-j)} &= \theta_{k,l}(N-j) + (\beta_{k,l} - \theta_{k,l}(N-j)) \frac{\tau_l EV(\mathbf{A} = [l, \dots, l, m])}{r_{(N-j),l}} \\ &+ \sum_{i=1}^{j-1} \left( \sum_{s=1}^i \left( \begin{array}{c} (\theta_{k,l}(N-j)\eta_l(N-j, N-j+s) \\ - \delta_{k,l}(N-j, N-j+s)) \\ \times \left( 1_{s=i} + \sum_{u=1}^{j-1-s} \left( \prod_{x=N-j+s}^{N-1-u} \eta_l(x, x+u) \right) \right) \right) \right) \\ &\times \left( \frac{r_{N-j+i,l} - \tau_l EV(\mathbf{A} = [l, \dots, l, m])}{r_{N-j,l}} \right) \end{array} \right) \quad (14) \end{aligned}$$

where

$$1_{s=1} = \begin{cases} 1 & \text{if } s = i, \\ 0 & \text{if } s \neq i, \end{cases}$$

is the indicator operator.

**Corollary 2.** There exists a set of critical ratios,  $\bar{\alpha}_{k,l}^i$  and  $\underline{\alpha}_{k,l}^i$  as expressed in Equations (13) and (14), respectively, that determines the firm's preference between products  $k$  and  $l$  in each state  $i = 1, \dots, N - 1$ . (i) In the case that  $EV(\mathbf{A} = [k, \dots, k, m]) > EV(\mathbf{A} = [l, \dots, l, m])$ , when  $RR_{k,l}^i > \bar{\alpha}_{k,l}^i$ , the firm prefers to manufacture product  $k$ , otherwise (when  $RR_{k,l}^i \leq \bar{\alpha}_{k,l}^i$ ), the firm prefers to manufacture product  $l$  in state  $i$ ; (ii) in the case that  $EV(\mathbf{A} = [k, \dots, k, m]) \leq EV(\mathbf{A} = [l, \dots, l, m])$ , when  $RR_{k,l}^i > \underline{\alpha}_{k,l}^i$ , the firm prefers to manufacture product  $k$ , otherwise (when  $RR_{k,l}^i \leq \underline{\alpha}_{k,l}^i$ ), the firm prefers to manufacture product  $l$  in state  $i$ .

The most general form of the critical ratios expressed in Equations (13) and (14) are not restricted by a functional relationship between the deterioration probabilities and the processing times. Although the above corollary establishes the preference relationships similar to those in Proposition 1, the relative values of the two critical ratios cannot be described uniformly as in Proposition 3 due to the lack of a functional relationship between processing times and deterioration probabilities. Thus, structural properties regarding the optimal solution similar to those presented in Theorem 1 cannot be characterized in this case. However, the firm can still make use of these critical ratios in order to determine its production preferences (and reduce the feasible set of potentially optimal policies) as they continue to pertain to the reservation prices.

## 5. Conclusions

This paper studies optimal production decisions for a single-stage manufacturing system where the system condition deteriorates over time, thus influencing the yield. The machine condition is characterized by a discrete number of states, and the goal of the decision-maker is to determine the optimal production choices in each state. We consider multiple products, and therefore the production decision corresponds to *which* product is optimal to be produced in each state. The decision to produce one product over the other impacts the machine condition, because expected processing times of these products are different, resulting in altered transition probabilities between states. The manufacturer performs maintenance when the machine worsens, and returns the equipment to its best state. As a result, a SMDP describes the process, and the steady-state probability for each state can be determined by using EMCs.

The set of decisions made in this paper and the corresponding analysis depart from earlier research on several levels. While traditional studies investigate the optimal production quantity (i.e., how much), our paper considers a multiple product environment, and thus, the decision corresponds to the *choice* of the product to be produced in each state. Owing to the complexity of the problem, earlier research commonly focuses on a smaller set of policies, such as monotone policies, and develops the sufficient conditions that make them optimal. In contrast, our technique is unique because it considers the entire set of potentially optimal policies, characterizes each one and develops the necessary and sufficient conditions to determine which one is optimal.

The paper makes three sets of contributions. The first set of contributions introduces the concept of *critical ratios* for the firm's decision at each state regarding which product to manufacture. These critical ratios, when multiplied by the reward of the reference product, provide the managerially insightful *reservation price* between two products. Put differently, it can be thought of as the least amount of money that a manager should be willing to earn in order to switch from producing one product to another.

The second set of contributions corresponds to the discovery that the optimal decision as to which product to produce in a state can be determined independently of the production decisions made in other states. This is a surprising and counter-intuitive result because the stationary probability corresponding to one state changes with the decisions made in all other states. Thus, one would not expect to develop a method that allows *separable* optimal decisions in each state. We prove that the optimal production decision in a particular state can be determined by comparing the ratio of rewards for two products with the critical ratio. These critical ratios are closed-form expressions that integrate the transition probabilities and the ratio of rewards with the varying processing time requirements of each product, including the mean and the variance.

The third set of contributions corresponds to the generalizations of these analytical results to more complex settings. When the problem is extended in the number of states that describe the machine condition, the firm has to compute only one additional set of critical ratios in order to determine the optimal production decision, rather than enumerating all of the potentially optimal policies in the new problem setting. In the next generalization, it is shown that when it is optimal to perform maintenance in an intermediate state, then it is optimal to maintain in all of the following worse (or higher) states. Thus, the problem can be reduced to a setting that includes only the states where production is preferred and the first state where maintenance is performed. The final generalization incorporates the impact of the variance in processing times on transition probabilities and shows how the critical ratios change with increasing variance. The solution approach prescribed in this paper is beneficial for further generalizations, including the case when the manufacturer has demand constraints. If the optimal solution obtained through our approach is feasible under demand constraints, then it is also optimal for the new problem setting. However, the optimal solution can be a single-product policy, and can be infeasible under demand constraints. When the optimal policy is infeasible under demand constraints, it can be assigned as the reference policy to determine a new set of critical ratios. In a two-product setting, only one of the demand constraints will be violated, and the new critical ratios can be used to obtain the least costly switches in order to determine the constrained optimal policy.

In sum, we show: (i) how a rich modeling framework (with a series of problem variants) can be developed for the important problem of production planning under deteriorating equipment condition; and (ii) the robustness of the critical ratios in the optimal solution algorithm.

## References

- 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. and Ord, J.K. (2000) The impact of imperfect processes on production run times. *Decision Sciences*, **31**, 773–787.
- Heyman, D.P. and Sobel, M.J. (1984) *Stochastic Models in Operations Research, Volume II: Stochastic Optimization*. McGraw-Hill, New York, NY.
- Howard, R.A. (1960) *Dynamic Programming and Markov Processes*. Technology Press of MIT, Cambridge, MA.
- Kao, E.P.C. (1973) Optimal replacement rules when changes of state are semi-Markovian. *Operations Research*, **21**, 1231–1249.
- Kim, C.H., Hong, Y.S. and Chang, S.Y. (2001) Optimal production run length and inspection schedules in a deteriorating production process. *IIE Transactions*, **33**, 421–426.
- Lee, H.L. and Rosenblatt, M.J. (1987) Simultaneous determination of production cycle and inspection schedules in a production system. *Management Science*, **33**, 1125–1136.

- Lee, H.L. and 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. and 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.
- Makis, V. and Fung, J. (1998) An EMQ model with inspections and random machine failures. *Journal of the Operational Research Society*, **49**, 66–76.
- 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, NY.
- Rosenblatt, M.J. and Lee, H.L. (1986) Economic production cycles with imperfect production processes. *IIE Transactions*, **18**, 48–54.
- Sloan, T.W. and 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. and 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.
- Tijms, H.C. (1986) *Stochastic Modelling and Analysis: A Computational Approach*, Wiley, New York, NY.
- Yano, C.A. and Lee, H.L. (1995) Lot sizing with random yields: a review. *Operations Research*, **43**, 311–334.
- Zequeira, R.I., Prida, B. and Valdes, J.E. (2004) Optimal buffer inventory and preventive maintenance for an imperfect production process. *International Journal of Production Research*, **42**, 959–974.

## Appendix

**Proof of Proposition 1.** We provide the proof for the case when  $EV(\mathbf{A}_4 = [2, 2, m]) \geq EV(\mathbf{A}_1 = [1, 1, m])$ , and thus,  $\alpha_{1,2}^1 = \alpha_{1,2}^2$  for each state  $i = 1, 2$  (the proof for the case when  $EV(\mathbf{A}_4 = [2, 2, m]) < EV(\mathbf{A}_1 = [1, 1, m])$  is similar). The critical ratio for state 2 can be obtained by equating  $EV(\mathbf{A}_3 = [2, 1, m]) = EV(\mathbf{A}_4 = [2, 2, m])$ :

$$\begin{aligned} \frac{r_{12}(1-p_{22}^1) + r_{21}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^1)}{\tau_2(1-p_{22}^1) + \tau_1p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^1)} &= \left( \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)} \right. \\ \left. \frac{c\beta_{1,2}r_{12}(1-p_{22}^2) + \alpha_{1,2}^2r_{22}p_{12}^2 + c\beta_{1,2}r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{c\beta_{1,2}\tau_2(1-p_{22}^2) + \beta_{1,2}\tau_2p_{12}^2 + c\beta_{1,2}\tau_m(1-p_{11}^2)(1-p_{22}^2)} \right) &= \left( \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)} \right. \\ \left. \frac{c\beta_{1,2}r_{12}(1-p_{22}^2) + \alpha_{1,2}^2r_{22}p_{12}^2 + c\beta_{1,2}r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{c\beta_{1,2}\tau_2(1-p_{22}^2) + c\beta_{1,2}\tau_2p_{12}^2 + c\beta_{1,2}\tau_m(1-p_{11}^2)(1-p_{22}^2) + (\beta_{1,2} - c\beta_{1,2})\tau_2p_{12}^2} \right) \\ &= \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)}, \\ \alpha_{1,2}^2r_{22}p_{12}^2 &= c\beta_{1,2}r_{22}p_{12}^2 + (\beta_{1,2} - c\beta_{1,2})\tau_2p_{12}^2EV(\mathbf{A}_4 = [2, 2, m]). \\ \alpha_{1,2}^2 &= c\beta_{1,2} + (\beta_{1,2} - c\beta_{1,2})\frac{\tau_2EV(\mathbf{A}_4 = [2, 2, m])}{r_{22}}. \end{aligned}$$

Similarly, the critical ratio for state 1 can be found by equating  $EV(\mathbf{A}_2 = [1, 2, m]) = EV(\mathbf{A}_4 = [2, 2, m])$ :

$$\begin{aligned} \frac{r_{11}(1-p_{22}^2) + r_{22}p_{12}^1 + r_{3m}(1-p_{11}^1)(1-p_{22}^2)}{\tau_1(1-p_{22}^2) + \tau_2p_{12}^1 + \tau_m(1-p_{11}^1)(1-p_{22}^2)} &= \left( \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)} \right. \\ \left. \frac{\alpha_{1,2}^1r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + c\beta_{1,2}r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\beta_{1,2}\tau_2(1-p_{22}^2) + c\beta_{1,2}\tau_2p_{12}^2 + c\beta_{1,2}\tau_m(1-p_{11}^2)(1-p_{22}^2)} \right) &= \left( \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)} \right. \\ \left. \frac{\alpha_{1,2}^1r_{12}(1-p_{22}^2) + c\beta_{1,2}r_{22}p_{12}^2 + c\beta_{1,2}r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{c\beta_{1,2}\tau_2(1-p_{22}^2) + c\beta_{1,2}\tau_2p_{12}^2 + c\beta_{1,2}\tau_m(1-p_{11}^2)(1-p_{22}^2) + (\beta_{1,2} - c\beta_{1,2})\tau_2(1-p_{22}^2)} \right) \\ &= \frac{r_{12}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)}. \\ \alpha_{1,2}^1r_{12}(1-p_{22}^2) &= c\beta_{1,2}r_{12}(1-p_{22}^2) + (\beta_{1,2} - c\beta_{1,2})\tau_2(1-p_{22}^2)EV(\mathbf{A}_4 = [2, 2, m]). \\ \alpha_{1,2}^1 &= c\beta_{1,2} + (\beta_{1,2} - c\beta_{1,2})\frac{\tau_2EV(\mathbf{A}_4 = [2, 2, m])}{r_{12}}. \end{aligned}$$

(i) State 1: The case when  $RR_{1,2}^1 > \underline{\alpha}_{1,2}^1$  is proven by substituting  $RR_{1,2}^1 r_{12}$  for  $r_{11}$  in  $EV(\mathbf{A}_2 = [1, 2, m])$ :

$$\begin{aligned} EV(\mathbf{A}_2 = [1, 2, m]) &= \frac{RR_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &> \frac{\underline{\alpha}_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

(i) State 2: The case when  $RR_{1,2}^2 > \underline{\alpha}_{1,2}^2$  is proven by substituting  $RR_{1,2}^2 r_{22}$  for  $r_{21}$  in  $EV(\mathbf{A}_3 = [2, 1, m])$ :

$$\begin{aligned} EV(\mathbf{A}_3 = [2, 1, m]) &= \frac{r_{12} (1 - p_{22}^1) + RR_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &> \frac{r_{12} (1 - p_{22}^1) + \underline{\alpha}_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

(ii) State 1: The case when  $RR_{1,2}^1 < \underline{\alpha}_{1,2}^1$  is proven by substituting  $RR_{1,2}^1 r_{12}$  for  $r_{11}$  in  $EV(\mathbf{A}_2 = [1, 2, m])$ :

$$\begin{aligned} EV(\mathbf{A}_2 = [1, 2, m]) &= \frac{RR_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &< \frac{\underline{\alpha}_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

(ii) State 2: The case when  $RR_{1,2}^2 < \underline{\alpha}_{1,2}^2$  is proven by substituting  $RR_{1,2}^2 r_{22}$  for  $r_{21}$  in  $EV(\mathbf{A}_3 = [2, 1, m])$ :

$$\begin{aligned} EV(\mathbf{A}_3 = [2, 1, m]) &= \frac{r_{12} (1 - p_{22}^1) + RR_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &< \frac{r_{12} (1 - p_{22}^1) + \underline{\alpha}_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

(iii) State 1: The case when  $RR_{1,2}^1 = \underline{\alpha}_{1,2}^1$  is proven by substituting  $RR_{1,2}^1 r_{12}$  for  $r_{11}$  in  $EV(\mathbf{A}_2 = [1, 2, m])$ :

$$\begin{aligned} EV(\mathbf{A}_2 = [1, 2, m]) &= \frac{RR_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &= \frac{\underline{\alpha}_{1,2}^1 r_{12} (1 - p_{22}^2) + r_{22} p_{12}^1 + r_{3m} (1 - p_{11}^1) (1 - p_{22}^2)}{\tau_1 (1 - p_{22}^2) + \tau_2 p_{12}^1 + \tau_m (1 - p_{11}^1) (1 - p_{22}^2)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

(iii) State 2: The case when  $RR_{1,2}^2 = \underline{\alpha}_{1,2}^2$  is proven by substituting  $RR_{1,2}^2 r_{22}$  for  $r_{21}$  in  $EV(\mathbf{A}_3 = [2, 1, m])$ :

$$\begin{aligned} EV(\mathbf{A}_3 = [2, 1, m]) &= \frac{r_{12} (1 - p_{22}^1) + RR_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &= \frac{r_{12} (1 - p_{22}^1) + \underline{\alpha}_{1,2}^2 r_{22} p_{12}^2 + r_{3m} (1 - p_{11}^2) (1 - p_{22}^1)}{\tau_2 (1 - p_{22}^1) + \tau_1 p_{12}^2 + \tau_m (1 - p_{11}^2) (1 - p_{22}^1)} \\ &= EV(\mathbf{A}_4 = [2, 2, m]). \end{aligned}$$

■

**Proof of Proposition 2.** In both of the critical ratio expressions, the first term  $c\beta_{1,2}$  is constant.

- (i) When  $c \geq 1$ , the second term in both critical ratio expressions, that is

$$\beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{i2}}$$

and

$$\beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{i2}},$$

are negative. Because  $r_{i2} \geq r_{j2}$ , where  $i < j$ , we get:

$$\begin{aligned} & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{i2}} \\ \leq & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{j2}}, \end{aligned}$$

and

$$\begin{aligned} & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{i2}} \\ \leq & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{j2}}. \end{aligned}$$

Therefore,  $\underline{\alpha}_{1,2}^i \geq \underline{\alpha}_{1,2}^j$ , and  $\bar{\alpha}_{1,2}^i \geq \bar{\alpha}_{1,2}^j$ . Thus,  $\underline{\alpha}_{1,2}^i$  and  $\bar{\alpha}_{1,2}^i$  are non-increasing in  $i$ , and therefore,  $\alpha_{1,2}^i$  is non-increasing in  $i$ .

- (ii) When  $\frac{1}{\beta_{1,2}} \leq c < 1$ , the second term in both critical ratio expressions, that is

$$\beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{i2}},$$

and

$$\beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{i2}},$$

are positive. Because  $r_{i2} \geq r_{j2}$ , where  $i < j$ , we get:

$$\begin{aligned} & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{i2}} \\ \geq & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_4 = [2, 2, m])}{r_{j2}}, \end{aligned}$$

and

$$\begin{aligned} & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{i2}} \\ \geq & \beta_{1,2}(1-c) \frac{\tau_2 EV(\mathbf{A}_1 = [1, 1, m])}{r_{j2}}. \end{aligned}$$

Therefore,  $\underline{\alpha}_{1,2}^i \leq \underline{\alpha}_{1,2}^j$ , and  $\bar{\alpha}_{1,2}^i \leq \bar{\alpha}_{1,2}^j$ . Thus,  $\underline{\alpha}_{1,2}^i$  and  $\bar{\alpha}_{1,2}^i$  are non-decreasing in  $i$ , and therefore,  $\alpha_{1,2}^i$  is non-decreasing in  $i$ . ■

**Proof of Proposition 3.**

- (i) When  $c \geq 1$ , we get  $\beta_{1,2}(1-c) < 0$ . Then, when  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ ,  $\underline{\alpha}_{1,2}^i - \bar{\alpha}_{1,2}^i$

$= \beta_{1,2}(1-c)\tau_2/r_{i2}[EV(\mathbf{A}_4 = [2, 2, m]) - EV(\mathbf{A}_1 = [1, 1, m])] > 0$ . Thus,  $\underline{\alpha}_{1,2}^i > \bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ .

- (ii) When  $c \geq 1$ , we get  $\beta_{1,2}(1-c) < 0$ . Then, when  $EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m])$ ,  $\underline{\alpha}_{1,2}^i - \bar{\alpha}_{1,2}^i = \beta_{1,2}(1-c)\tau_2/r_{i2}[EV(\mathbf{A}_4 = [2, 2, m]) - EV(\mathbf{A}_1 = [1, 1, m])] \leq 0$ . Thus,  $\underline{\alpha}_{1,2}^i \leq \bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ .

- (iii) When  $1/\beta_{1,2} \leq c < 1$ , we get  $\beta_{1,2}(1-c) > 0$ . Then, when  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ ,  $\underline{\alpha}_{1,2}^i - \bar{\alpha}_{1,2}^i = \beta_{1,2}(1-c)\tau_2/r_{i2}[EV(\mathbf{A}_4 = [2, 2, m]) - EV(\mathbf{A}_1 = [1, 1, m])] < 0$ . Thus,  $\underline{\alpha}_{1,2}^i < \bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ .

- (iv) When  $1/\beta_{1,2} \leq c < 1$ , we get  $\beta_{1,2}(1-c) > 0$ . Then, when  $EV(\mathbf{A}_1 = [1, 1, m]) \leq EV(\mathbf{A}_4 = [2, 2, m])$ ,  $\underline{\alpha}_{1,2}^i - \bar{\alpha}_{1,2}^i = \beta_{1,2}(1-c)\tau_2/r_{i2}[EV(\mathbf{A}_4 = [2, 2, m]) - EV(\mathbf{A}_1 = [1, 1, m])] \geq 0$ . Thus,  $\underline{\alpha}_{1,2}^i \geq \bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ . ■

**Proof of Theorem 1.** We provide the proof for the case when  $c \geq 1$ , and the proof for the case when  $1/\beta_{1,2} \leq c < 1$  is similar.

1. The case when  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$  and thus  $\alpha_{1,2}^i = \bar{\alpha}_{1,2}^i$  for each  $i = 1, 2$ .

Under these conditions, we already know from Proposition 3(i) that  $\underline{\alpha}_{1,2}^i \geq \bar{\alpha}_{1,2}^i$  for each  $i = 1, 2$ . The value of the ratio of rewards can be in one of four possible scenarios.

*Scenario 1:*  $\bar{\alpha}_{1,2}^1 \leq RR_{1,2}^1$  and  $\bar{\alpha}_{1,2}^2 \leq RR_{1,2}^2$ .

From Proposition 1, we know that  $\bar{\alpha}_{1,2}^1 \leq RR_{1,2}^1$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_3 = [2, 1, m])$  and  $\bar{\alpha}_{1,2}^2 \leq RR_{1,2}^2$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_2 = [1, 2, m])$ . By the definition of this case, we already have  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ . Therefore,  $EV(\mathbf{A}_1 = [1, 1, m])$  is the highest expected reward collectively, and producing product 1 is optimal in both states,  $i = 1, 2$  (i.e.,  $a_1^* = 1, a_2^* = 1$ ).

*Scenario 2:*  $\bar{\alpha}_{1,2}^1 \leq RR_{1,2}^1$  and  $\bar{\alpha}_{1,2}^2 > RR_{1,2}^2$ .

From Proposition 1, we know that  $\bar{\alpha}_{1,2}^1 \leq RR_{1,2}^1$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_3 = [2, 1, m])$ , and  $\bar{\alpha}_{1,2}^2 > RR_{1,2}^2$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_2 = [1, 2, m])$ . By the definition of this case, we already have  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ . Therefore, the expected values are in the following order:

$$\begin{aligned} & EV(\mathbf{A}_2 = [1, 2, m]) > EV(\mathbf{A}_1 = [1, 1, m]) \\ \geq & \begin{cases} EV(\mathbf{A}_3 = [2, 1, m]) \\ EV(\mathbf{A}_4 = [2, 2, m]) \end{cases}. \end{aligned}$$

Thus, policy  $\mathbf{A}_2 = [1, 2, m]$  is optimal. This leads to the optimal choices of product 1 in state 1 and product 2 in state 2, i.e.,  $a_1^* = 1, a_2^* = 2$ .

*Scenario 3:*  $\bar{\alpha}_{1,2}^1 > RR_{1,2}^1$  and  $\bar{\alpha}_{1,2}^2 \leq RR_{1,2}^2$ .

From Proposition 1, we know that  $\bar{\alpha}_{12}^1 > RR_{1,2}^1$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_3 = [2, 1, m])$  and  $\bar{\alpha}_{12}^2 \leq RR_{1,2}^2$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_2 = [1, 2, m])$ . By the definition of this case, we already have  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ . Therefore, the expected values are in the following order:

$$\begin{aligned} & EV(\mathbf{A}_3 = [2, 1, m]) > EV(\mathbf{A}_1 = [1, 1, m]) \\ & \geq \left\{ \begin{array}{l} EV(\mathbf{A}_2 = [1, 2, m]) \\ EV(\mathbf{A}_4 = [2, 2, m]) \end{array} \right\}. \end{aligned}$$

Thus, policy  $EV(\mathbf{A}_3 = [2, 1, m])$  is optimal. This leads to the optimal choices of product 2 in state 1 and product 1 in state 2, i.e.,  $a_1^* = 2, a_2^* = 1$ .

*Scenario 4:*  $\bar{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\bar{\alpha}_{12}^2 > RR_{1,2}^2$ .

From Proposition 1, we know that  $\bar{\alpha}_{12}^1 > RR_{1,2}^1$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_3 = [2, 1, m])$  and  $\bar{\alpha}_{12}^2 > RR_{1,2}^2$  implies that  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_2 = [1, 2, m])$ . Thus,

$$\left\{ \begin{array}{l} EV(\mathbf{A}_2 = [1, 2, m]) \\ EV(\mathbf{A}_3 = [2, 1, m]) \end{array} \right\} > EV(\mathbf{A}_1 = [1, 1, m]).$$

However, we already know from Proposition 3(i) that  $\underline{\alpha}_{1,2}^i \geq \bar{\alpha}_{12}^i$  for each  $i = 1, 2$ . Therefore in this scenario, we have  $\underline{\alpha}_{1,2}^1 \geq \bar{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\underline{\alpha}_{1,2}^2 \geq \bar{\alpha}_{12}^2 > RR_{1,2}^2$ . From Proposition 1, we know that  $\underline{\alpha}_{1,2}^1 > RR_{1,2}^1$  implies  $EV(\mathbf{A}_4 = [2, 2, m]) > EV(\mathbf{A}_2 = [1, 2, m])$  and that  $\underline{\alpha}_{1,2}^2 > RR_{1,2}^2$  implies  $EV(\mathbf{A}_4 = [2, 2, m]) > EV(\mathbf{A}_3 = [2, 1, m])$ . Collectively, we get:

$$EV(\mathbf{A}_4 = [2, 2, m]) > \left\{ \begin{array}{l} EV(\mathbf{A}_2 = [1, 2, m]) \\ EV(\mathbf{A}_3 = [2, 1, m]) \end{array} \right\}.$$

When these are combined with the earlier comparisons, we get:

$$\begin{aligned} EV(\mathbf{A}_4 = [2, 2, m]) & > \left\{ \begin{array}{l} EV(\mathbf{A}_2 = [1, 2, m]) \\ EV(\mathbf{A}_3 = [2, 1, m]) \end{array} \right\} \\ & > EV(\mathbf{A}_1 = [1, 1, m]), \end{aligned}$$

contradicting the motivating case of  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ . As a result, this scenario is never encountered when  $EV(\mathbf{A}_1 = [1, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ , proving part (iii) of the theorem. Scenarios 1, 2 and 3 collectively prove that when  $\bar{\alpha}_{12}^i \leq RR_{1,2}^i$ , then the optimal production decision is  $a_i^* = 1$ , and when  $\bar{\alpha}_{12}^i > RR_{1,2}^i$ , then the optimal production decision is  $a_i^* = 2$ . This completes parts (i) and (ii) of the proof of the theorem.

2. The case when  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_4 = [2, 2, m])$ .

Under these conditions, we already know from Proposition 3(ii) that  $\underline{\alpha}_{1,2}^i < \bar{\alpha}_{12}^i$  for each  $i = 1, 2$ . The value of the ratio of rewards can be in one of four possible scenarios.

*Scenario 1:*  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$ .

From Proposition 1, we know that  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  implies that  $EV(\mathbf{A}_2 = [1, 2, m]) < EV(\mathbf{A}_4 = [2, 2, m])$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$  implies that  $EV(\mathbf{A}_3 = [2, 1, m]) < EV(\mathbf{A}_4 = [2, 2, m])$ . As a result, we have:

$$EV(\mathbf{A}_4 = [2, 2, m]) > \left\{ \begin{array}{l} EV(\mathbf{A}_1 = [1, 1, m]) \\ EV(\mathbf{A}_2 = [1, 2, m]) \\ EV(\mathbf{A}_3 = [2, 1, m]) \end{array} \right\}.$$

Thus, the optimal policy is  $\mathbf{A}_4 = [2, 2, m]$  and the optimal production choice is product 2 in both states, i.e.,  $a_1^* = 2, a_2^* = 2$ .

*Scenario 2:*  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\underline{\alpha}_{12}^2 \leq RR_{1,2}^2$ .

From Proposition 1, we know that  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  implies that  $EV(\mathbf{A}_2 = [1, 2, m]) < EV(\mathbf{A}_4 = [2, 2, m])$  and  $\underline{\alpha}_{12}^2 \leq RR_{1,2}^2$  implies that  $EV(\mathbf{A}_3 = [2, 1, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$ . By the definition of this case, we already have  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_4 = [2, 2, m])$ . Therefore, the expected values are in the following order:

$$\begin{aligned} EV(\mathbf{A}_3 = [2, 1, m]) & \geq EV(\mathbf{A}_4 = [2, 2, m]) \\ & > \left\{ \begin{array}{l} EV(\mathbf{A}_1 = [1, 1, m]) \\ EV(\mathbf{A}_2 = [1, 2, m]) \end{array} \right\}. \end{aligned}$$

Thus, policy  $\mathbf{A}_3 = [2, 1, m]$  is optimal. This leads to the optimal choices of product 2 in state 1 and product 1 in state 2, i.e.,  $a_1^* = 2, a_2^* = 1$ .

*Scenario 3:*  $\underline{\alpha}_{12}^1 \leq RR_{1,2}^1$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$ :

From Proposition 1, we know that  $\underline{\alpha}_{12}^1 \leq RR_{1,2}^1$  implies that  $EV(\mathbf{A}_2 = [1, 2, m]) \geq EV(\mathbf{A}_4 = [2, 2, m])$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$  implies that  $EV(\mathbf{A}_3 = [2, 1, m]) < EV(\mathbf{A}_4 = [2, 2, m])$ . By the definition of this case, we already have  $EV(\mathbf{A}_1 = [1, 1, m]) < EV(\mathbf{A}_4 = [2, 2, m])$ . Therefore, the expected values are in the following order:

$$\begin{aligned} EV(\mathbf{A}_2 = [1, 2, m]) & \geq EV(\mathbf{A}_4 = [2, 2, m]) \\ & > \left\{ \begin{array}{l} EV(\mathbf{A}_1 = [1, 1, m]) \\ EV(\mathbf{A}_3 = [2, 1, m]) \end{array} \right\}. \end{aligned}$$

Thus, policy  $\mathbf{A}_2 = [1, 2, m]$  is optimal. This leads to the optimal choices of product 1 in state 1 and product 2 in state 2, i.e.,  $a_1^* = 1, a_2^* = 2$ .

*Scenario 4:*  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$ .

It can easily be seen that when both  $\underline{\alpha}_{12}^1 > RR_{1,2}^1$  and  $\underline{\alpha}_{12}^2 > RR_{1,2}^2$ , we get  $EV(\mathbf{A}_1 = [1, 1, m]) > EV(\mathbf{A}_4 = [2, 2, m])$ , and this contradicts the original case, proving part (iii) of the theorem.

Scenarios 1, 2 and 3 collectively prove that when  $\underline{\alpha}_{12}^i \leq RR_{1,2}^i$ , then the optimal production decision is  $a_i^* = 1$ , and when  $\underline{\alpha}_{12}^i > RR_{1,2}^i$ , then the optimal production decision is  $a_i^* = 2$ . This completes parts (i) and (ii) of the proof of the theorem. ■

#### Proof of Proposition 4.

The proof follows from Propositions 2 and 3.



- (i) When  $c > 1$ , both critical ratios are non-increasing in  $i$ . Therefore, in the case of constant ratio of rewards, i.e.,  $RR_{1,2}^1 = RR_{1,2}^2 = RR_{1,2}^i = RR_{1,2}$  for all  $i = 1, \dots, N$ , we get  $RR_{1,2} - \underline{\alpha}_{1,2}^i \leq RR_{1,2} - \underline{\alpha}_{1,2}^j$  for all states  $i < j$  as well as  $RR_{1,2} - \bar{\alpha}_{1,2}^i \leq RR_{1,2} - \bar{\alpha}_{1,2}^j$  for all states  $i < j$ ; and, their signs can switch from negative to positive only once.
- (ii) When  $1/\beta_{1,2} \leq c < 1$ , both critical ratios are non-decreasing in  $i$ . Therefore, in the case of constant ratio of rewards, i.e.,  $RR_{1,2}^1 = RR_{1,2}^2 = RR_{1,2}^i = RR_{1,2}$ , we get  $RR_{1,2} - \underline{\alpha}_{1,2}^i \geq RR_{1,2} - \underline{\alpha}_{1,2}^j$  for all states  $i < j$  as well as  $RR_{1,2} - \bar{\alpha}_{1,2}^i \geq RR_{1,2} - \bar{\alpha}_{1,2}^j$  for all states  $i < j$ ; and, their signs can switch from positive to negative only once. ■

**Proof of Corollary 1.** The proof follows directly from Theorem 1. ■

**Proof of Lemma 1.** The proof follows from Theorem 3 of Kao (1973) and the fact that the reward structure is non-increasing in the machine state, i.e.,  $r_{ia} \geq r_{ja}$  for each action  $a$  and for all states  $1 \leq i < j \leq N$ . ■

**Proof of Proposition 5.** When maintenance is performed, the machine returns to its best state with probability  $p_{i1}^m = 1$  for all  $i = 2, \dots, N$ . First, consider the problem setting with  $N$  states. The steady-state probability for states where production takes place, i.e.,  $i = 1, \dots, M-1$ , using any policy  $\mathbf{A}_N = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m, a_N = m]$  generates  $\pi_i(\mathbf{A}_N = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m, a_N = m])$ . Now, consider the problem setting with  $N-1$  states, and a policy that uses the same sequence of actions in all states from  $i = 1$  to  $i = N-1$ , denoted by  $\mathbf{A}_{N-1}$ . The steady-state probability for this policy in the states that production takes place, i.e.,  $i = 1, \dots, M-1$ , is denoted by  $\pi_i(\mathbf{A}_{N-1} = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m])$ . It should be observed that  $\pi_i(\mathbf{A}_N = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m, a_N = m]) = \pi_i(\mathbf{A}_{N-1} = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m]) \times p_{N1}^m$ . Because  $p_{N1}^m = 1$ , we get  $\pi_i(\mathbf{A}_N = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m, a_N = m]) = \pi_i(\mathbf{A}_{N-1} = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m])$ . Therefore, the expected values of these two policies, despite the fact that they are in two different settings, are equal. Thus,  $EV(\mathbf{A}_N) = EV(\mathbf{A}_{N-1})$ . By induction, we get  $\pi_i(\mathbf{A}_N = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m, a_N = m]) = \pi_i(\mathbf{A}_{N-1} = [a_1, \dots, a_{M-1}, a_M = m, \dots, a_{N-1} = m]) = \dots = \pi_i(\mathbf{A}_M = [a_1, \dots, a_{M-1}, a_M = m])$ . Moreover, we get  $EV(\mathbf{A}_N) = EV(\mathbf{A}_{N-1}) = \dots = EV(\mathbf{A}_M)$ . Let us denote  $\mathbf{A}_N(1)$  and  $\mathbf{A}_N(2)$  as the single-product policies of manufacturing products 1 and 2, respectively, in a problem setting that features  $N$  machine states. We next show that the critical ratios for problem settings that have production actions in the first  $M-1$  states and maintenance in

the following (worse) states are equal for all problem settings that feature  $M$  or more machine states. Thus,  $\underline{\alpha}_{k,l}^i(N, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_N(2)))/r_{i2} = \underline{\alpha}_{k,l}^i(N-1, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_{N-1}(2)))/r_{i2} = \dots = \underline{\alpha}_{k,l}^i(M, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_M(2)))/r_{i2}$  for all  $i = 1, \dots, M-1$ , and  $\bar{\alpha}_{k,l}^i(N, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_N(1)))/r_{i2} = \bar{\alpha}_{k,l}^i(N-1, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_{N-1}(1)))/r_{i2} = \dots = \bar{\alpha}_{k,l}^i(M, M) = c\beta_{1,2} + \beta_{1,2}(1-c)(\tau_2 EV(\mathbf{A}_M(1)))/r_{i2}$  for all  $i = 1, \dots, M-1$ . As a result, we have  $\alpha_{k,l}^i(N, M) = \alpha_{k,l}^i(N-1, M) = \dots = \alpha_{k,l}^i(M, M)$  for all states  $i = 1, \dots, M-1$ . ■

**Proof of Proposition 6.** Products 1 and 2 have the same mean processing time, however, the variance of processing time is higher for product 1 than for product 2. Therefore, we have  $\tau_1 = \tau_2$ , and  $\sigma_1^2 > \sigma_2^2$ . We provide the proof for the case when  $\alpha_{1,2}^1 = \bar{\alpha}_{1,2}^1$ , and the proof for the case when  $\alpha_{1,2}^1 = \underline{\alpha}_{1,2}^1$  is similar. Let us determine the critical ratios defined by  $\bar{\alpha}_{1,2}^i$  for each state  $i = 1, 2$ .

(i) For state 1, we have  $EV[1, 1, m] = EV[2, 1, m]$ :

$$\begin{aligned} & \frac{r_{11}(1-p_{22}^1) + r_{21}p_{12}^1 + r_{3m}(1-p_{11}^1)(1-p_{22}^1)}{\tau_2(1-p_{22}^1) + \tau_2p_{12}^1 + \tau_m(1-p_{11}^1)(1-p_{22}^1)} \\ &= \frac{r_{12}(1-p_{22}^1) + r_{21}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^1)}{\tau_2(1-p_{22}^1) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^1)}. \end{aligned}$$

Substituting  $p_{12}^1 = p_{12}^2 + \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)$  and  $(1-p_{11}^1) = (1-p_{11}^2) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)$  into  $EV[2, 1, m]$  provides  $r_{11}(1-p_{22}^1) = r_{12}(1-p_{22}^1) - \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)(r_{21} - \tau_2 EV[1, 1, m]) + \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1-p_{22}^1)(r_{3m} - \tau_m EV[1, 1, m])$ . Thus,

$$\begin{aligned} \frac{r_{11}}{r_{12}} &= \bar{\alpha}_{1,2}^1 \\ &= 1 - \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{12}(1-p_{22}^1)}(r_{21} - \tau_2 EV[1, 1, m]) \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{12}}(r_{3m} - \tau_m EV[1, 1, m]). \end{aligned}$$

The critical ratio is greater than one when  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1-p_{22}^1)(r_{3m} - \tau_m EV[1, 1, m]) > \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)(r_{21} - \tau_2 EV[1, 1, m])$ . Therefore when this condition is satisfied,  $\bar{\alpha}_{1,2}^1$  is increasing in variance; otherwise, it is decreasing in the variance of processing times.

(ii) Similarly, for state 2, we have  $EV[1, 1, m] = EV[1, 2, m]$ :

$$\begin{aligned} & \frac{r_{11}(1-p_{22}^1) + r_{21}p_{12}^1 + r_{3m}(1-p_{11}^1)(1-p_{22}^1)}{\tau_2(1-p_{22}^1) + \tau_2p_{12}^1 + \tau_m(1-p_{11}^1)(1-p_{22}^1)} \\ &= \frac{r_{11}(1-p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1-p_{11}^2)(1-p_{22}^2)}{\tau_2(1-p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1-p_{11}^2)(1-p_{22}^2)}. \end{aligned}$$

Substituting  $(1-p_{11}^1) = (1-p_{11}^2) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)$  into  $EV[1, 2, m]$  provides  $r_{21}p_{12}^1 = r_{22}p_{12}^2 + \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(r_{21} - \tau_2 EV[1, 1, m]) + \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1-p_{11}^2)(r_{3m} - \tau_m EV[1, 1, m])$ .

Thus,

$$\begin{aligned} \frac{r_{21}}{r_{22}} &= \bar{\alpha}_{1,2}^2 \\ &= 1 + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{22}p_{12}^1} (r_{11} - \tau_2 EV[1, 1, m]) \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{22}p_{12}^1} (1 - p_{11}^1) (r_{3m} - \tau_m EV[1, 1, m]). \end{aligned}$$

The critical ratio is greater than one when  $r_{11} - \tau_2 EV[1, 1, m] < -(1 - p_{11}^1)(r_{3m} - \tau_m EV[1, 1, m])$ . Therefore, when this condition is satisfied,  $\bar{\alpha}_{1,2}^2$  is increasing in variance; otherwise, it is decreasing in the variance of processing times. ■

**Proof of Proposition 7.** The critical ratios for the comparison of products 2 and 3 were developed in Proposition 1. The comparison of products 1 and 3 captures the combined effect of mean and variance in processing times. Product 1 has a higher mean and variance in its processing time than product 3. Thus,  $\tau_1 = \tau_2 > \tau_3$  and  $\sigma_1^2 > \sigma_2^2 = \sigma_3^2$ . We first develop the critical ratios defined by  $\underline{\alpha}_{i,3}^1$  for each state  $i = 1, 2$ .

- (i) We can determine the critical ratio for state 1 by equating  $EV[1, 3, m] = EV[3, 3, m]$ :

$$\begin{aligned} \frac{r_{11}(1 - p_{22}^3) + r_{21}p_{12}^1 + r_{3m}(1 - p_{11}^1)(1 - p_{22}^3)}{\tau_2(1 - p_{22}^3) + \tau_3p_{12}^1 + \tau_m(1 - p_{11}^1)(1 - p_{22}^3)} \\ = \frac{r_{13}(1 - p_{22}^3) + r_{23}p_{12}^3 + r_{3m}(1 - p_{11}^3)(1 - p_{22}^3)}{\tau_3(1 - p_{22}^3) + \tau_3p_{12}^3 + \tau_m(1 - p_{11}^3)(1 - p_{22}^3)}. \end{aligned}$$

Note that  $p_{12}^1 = c\beta_{2,3}p_{12}^3 + \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)$  and  $(1 - p_{11}^1)(1 - p_{22}^3) = c\beta_{2,3}(1 - p_{11}^3)(1 - p_{22}^3) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{11}^3)$ . Substituting these two expressions in  $EV[1, 3, m]$  provides.

$$\begin{aligned} \frac{r_{11}}{r_{13}} &= \underline{\alpha}_{1,3}^1 = c\beta_{2,3} - (c\beta_{2,3} - \beta_{2,3}) \frac{\tau_3 EV[3, 3, m]}{r_{13}} \\ &\quad - \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1 - p_{22}^3)} (r_{23} - \tau_3 EV[3, 3, m]) \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}} (r_{3m} - \tau_m EV[3, 3, m]). \end{aligned}$$

Thus

$$\begin{aligned} \underline{\alpha}_{1,3}^1 &= \underline{\alpha}_{2,3}^1 - \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1 - p_{22}^3)} (r_{23} - \tau_3 EV[3, 3, m]) \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}} (r_{3m} - \tau_m EV[3, 3, m]). \end{aligned}$$

The critical ratio  $\underline{\alpha}_{1,3}^1$  is greater than  $\underline{\alpha}_{2,3}^1$  when  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{22}^3)(r_{3m} - \tau_m EV[3, 3, m]) > \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)(r_{23} - \tau_3 EV[3, 3, m])$ . Therefore, when this condition is satisfied, the firm needs to earn a higher reward to switch from product 3 to product 1 than from product 3 to product 2 in state 1.

- (ii) Similarly, we can obtain the critical ratio for state 2 by equating  $EV[3, 1, m] = EV[3, 3, m]$ :

$$\begin{aligned} \frac{r_{13}(1 - p_{22}^1) + r_{21}p_{12}^3 + r_{3m}(1 - p_{11}^3)(1 - p_{22}^1)}{\tau_3(1 - p_{22}^1) + \tau_2p_{12}^3 + \tau_m(1 - p_{11}^3)(1 - p_{22}^1)} \\ = \frac{r_{13}(1 - p_{22}^3) + r_{23}p_{12}^3 + r_{3m}(1 - p_{11}^3)(1 - p_{22}^3)}{\tau_3(1 - p_{22}^3) + \tau_3p_{12}^3 + \tau_m(1 - p_{11}^3)(1 - p_{22}^3)}. \end{aligned}$$

Note that  $(1 - p_{22}^1) = c\beta_{2,3}(1 - p_{22}^3) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)$  and  $(1 - p_{11}^1)(1 - p_{22}^3) = c\beta_{2,3}(1 - p_{11}^3)(1 - p_{22}^3) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{11}^3)$ . Substituting these two expressions in  $EV[3, 1, m]$  provides:

$$\begin{aligned} \frac{r_{21}}{r_{23}} &= \underline{\alpha}_{1,3}^2 = c\beta_{2,3} - (c\beta_{2,3} - \beta_{2,3}) \frac{\tau_3 EV[3, 3, m]}{r_{23}} \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^3} (r_{13} - \tau_3 EV[3, 3, m]) \\ &\quad + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^3} (1 - p_{11}^3)(r_{3m} - \tau_m EV[3, 3, m]). \end{aligned}$$

Thus,

$$\begin{aligned} \underline{\alpha}_{1,3}^2 &= \underline{\alpha}_{2,3}^2 + \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^3} [(r_{13} - \tau_3 EV[3, 3, m]) \\ &\quad + (1 - p_{11}^3)(r_{3m} - \tau_m EV[3, 3, m])]. \end{aligned}$$

The critical ratio  $\underline{\alpha}_{1,3}^2$  is greater than  $\underline{\alpha}_{2,3}^2$  when  $r_{13} - \tau_3 EV[3, 3, m] < -(1 - p_{11}^3)(r_{3m} - \tau_m EV[3, 3, m])$ . Therefore, when this condition is satisfied, the firm needs to earn a higher reward to switch from product 3 to product 1 than from product 3 to product 2 in state 2.

Using  $EV[2, 2, m]$  as the calibrating reference policy, we next develop the critical ratios  $\bar{\alpha}_{i,3}^1$  for each state  $i = 1, 2$ .

- (iii) For state 1,  $EV[3, 1, m] = EV[2, 2, m]$ :

$$\begin{aligned} \frac{r_{13}(1 - p_{22}^1) + r_{21}p_{12}^3 + r_{3m}(1 - p_{11}^3)(1 - p_{22}^1)}{\tau_3(1 - p_{22}^1) + \tau_2p_{12}^3 + \tau_m(1 - p_{11}^3)(1 - p_{22}^1)} \\ = \frac{r_{12}(1 - p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1 - p_{11}^2)(1 - p_{22}^2)}{\tau_2(1 - p_{22}^2) + \tau_2p_{12}^2 + \tau_m(1 - p_{11}^2)(1 - p_{22}^2)}. \end{aligned}$$

Note that  $(1 - p_{22}^1) = (1 - p_{22}^2) - \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)$ , and  $p_{12}^3 = (1/c\beta_{2,3}p_{12}^2)$ , and

$$\begin{aligned} (1 - p_{11}^3)(1 - p_{22}^1) &= \frac{1}{c\beta_{2,3}} (1 - p_{11}^2)(1 - p_{22}^2) \\ &\quad - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{11}^2)}{c\beta_{2,3}}. \end{aligned}$$

Substituting these three expressions in  $EV[3, 1, m]$  provides:

$$\begin{aligned} \bar{\alpha}_{1,3}^1 &= c\beta_{2,3} - (c\beta_{2,3} - \beta_{2,3}) \frac{\tau_2 EV[2, 2, m]}{r_{13}} \\ &\quad - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1 - p_{11}^2)} c\beta_{2,3} (r_{13} - \tau_2 EV[2, 2, m]) \\ &\quad - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}} (r_{3m} - \tau_m EV[2, 2, m]). \end{aligned}$$

Thus,

$$\bar{\alpha}_{13}^1 = \bar{\alpha}_{23}^1 - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)}{r_{13}(1 - p_{11}^2)} [c\beta_{23}(r_{13} - \tau_2 EV[2, 2, m]) + (1 - p_{11}^2)(r_{3m} - \tau_m EV[2, 2, m])].$$

Because  $\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2) < 0$ , the critical ratio  $\bar{\alpha}_{1,3}^1$  is greater than  $\bar{\alpha}_{2,3}^1$  when  $c\beta_{23}(r_{13} - \tau_2 EV[2, 2, m]) > -(1 - p_{22}^2)(r_{3m} - \tau_m EV[2, 2, m])$ . Therefore, when this condition is satisfied, the firm needs to earn a higher reward to switch from product 1 to product 3 than from product 2 to 3 in state 1.

(iv) For state 2,  $EV[1, 3, m] = EV[2, 2, m]$ :

$$\begin{aligned} & \frac{r_{11}(1 - p_{22}^3) + r_{23}p_{12}^1 + r_{3m}(1 - p_{11}^1)(1 - p_{22}^3)}{\tau_2(1 - p_{22}^3) + \tau_3 p_{12}^1 + \tau_m(1 - p_{11}^1)(1 - p_{22}^3)} \\ &= \frac{r_{12}(1 - p_{22}^2) + r_{22}p_{12}^2 + r_{3m}(1 - p_{11}^2)(1 - p_{22}^2)}{\tau_2(1 - p_{22}^2) + \tau_2 p_{12}^2 + \tau_m(1 - p_{11}^2)(1 - p_{22}^2)}. \end{aligned}$$

Note that  $(1 - p_{22}^3) = (1/c\beta_{2,3})(1 - p_{22}^2)$  and  $p_{12}^1 = p_{12}^2 + \varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)$  and

$$(1 - p_{11}^1)(1 - p_{22}^3) = \frac{1}{c\beta_{2,3}}(1 - p_{11}^2)(1 - p_{22}^2) - \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{22}^2)}{c\beta_{2,3}}.$$

Substituting these three expressions in  $EV[1, 3, m]$  provides:

$$\begin{aligned} \bar{\alpha}_{1,3}^2 &= c\beta_{2,3} - (c\beta_{2,3} - \beta_{2,3}) \frac{\tau_2 EV[2, 2, m]}{r_{23}} \\ &+ \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^2} c\beta_{23}(r_{23} - \tau_2 EV[2, 2, m]) \\ &- \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{11}^2)}{r_{23}p_{12}^3} (r_{3m} - \tau_m EV[2, 2, m]). \end{aligned}$$

Thus,

$$\begin{aligned} \bar{\alpha}_{13}^2 &= \bar{\alpha}_{23}^2 + \frac{\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)}{r_{23}p_{12}^2} c\beta_{23}(r_{23} - \tau_2 EV[2, 2, m]) \\ &- \frac{\varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(1 - p_{11}^2)}{r_{23}p_{12}^3} (r_{3m} - \tau_m EV[2, 2, m]). \end{aligned}$$

The critical ratio  $\bar{\alpha}_{1,3}^2$  is greater than  $\bar{\alpha}_{2,3}^2$  when  $\varepsilon_{12}^{12}(\sigma_1^2, \sigma_2^2)c\beta_{23}(r_{23} - \tau_2 EV[2, 2, m]) > \varepsilon_{11}^{12}(\sigma_1^2, \sigma_2^2)(r_{3m} - \tau_m EV[2, 2, m])$ . Therefore, when this condition is satisfied, the firm needs to earn a higher reward to switch from product 1 to product 3 than from product 2 to 3 in state 2. ■

**Proof of Corollary 2.** The proof follows directly from Corollary 1. ■

### Biographies

Burak Kazaz is an Assistant Professor of Supply Chain Management at Syracuse University. His current research interests include pricing and production planning problems that investigate the interactions between operations, finance and marketing. He earned his Ph.D. from the Krannert Graduate School of Management at Purdue University, and a B.S. and M.S. in Industrial Engineering from the Middle East Technical University in Ankara, Turkey. His research has appeared in academic journals such as *Management Science*, *Manufacturing & Service Operations Management* and the *European Journal of Operational Research*. He previously served as a School of Business Faculty at the University of Miami and the Loyola University of Chicago.

Thomas Sloan is an Assistant Professor at the University of Massachusetts Lowell. He received a B.B.A. degree from the University of Texas at Austin and M.S. and Ph.D. degrees from the University of California, Berkeley. His current research focuses on sustainable operations, particularly in the area of medical devices. He has also conducted research on shop-floor control and production scheduling in the semiconductor industry. His work has been published in academic journals such as the *International Journal of Production Research*, *Production and Operations Management*, *IEEE Transactions on Semiconductor Manufacturing*, *Journal of the Operational Research Society*, *IIE Transactions* and *Health Care Management Science*.