DYNAMIC SCHEDULING IN A MAKE-TO-STOCK SYSTEM: A PARTIAL CHARACTERISATION OF OPTIMAL POLICIES

Francis de Véricourt
Fikri Karaesmen
and
Yves Dallery

Laboratoire d'Informatique de Paris 6 (LIP6-CNRS)
Université Pierre et Marie Curie
4, place Jussieu
75252 Paris Cedex 05, France
Francis.Vericourt@lip6.fr, Fikri.Karaesmen@lip6.fr, Yves.Dallery@lip6.fr

DYNAMIC SCHEDULING IN A MAKE-TO-STOCK SYSTEM: A PARTIAL CHARACTERISATION OF OPTIMAL POLICIES

FRANCIS DE VERICOURT, FIKRI KARAESMEN
YVES DALLERY
Laboratoire d'Informatique de Paris 6 (LIP6-CNRS)
Universite Pierre et Marie Curie

May 28, 1998

Key Words: Production/scheduling, Stochastic: Multi-item; Queues: make-to-stock system; Inventory/production: Optimal policies.

Abstract

We consider the problem of dynamically allocating production capacity between two products to minimize the average inventory and backorder costs per unit time in a make-to-stock single machine system. Using sample path comparisons and dynamic programming, we give a characterization of the optimal hedging point policy for a certain region of the state space. The characterization is simple enough to lead to easily implementable heuristics and provides a formal justification of some of the earlier heuristics proposed.

1 Introduction

A challenging problem in production control is the dynamic allocation of limited production capacity between different products in a make-to-stock environment. The fact that demands and even the production times are random makes this problem even more challenging. In this paper, we provide new insights for the dynamic scheduling problem of a stochastic production-inventory system.

The particular model that we consider here is a two part-type model where demands of both types arrive in single units and a single production facility produces units one by one. The model is, then, the make-to-stock version of the well known multi-class single server queue, i.e. a two class *make-to-stock queue*. The question is to decide dynamically when and which part type to produce. For tractability, we make the usual assumptions that the demands arrive according to independent Poisson processes and that the production times are exponentially distributed. We also disregard setup times and allow preemptive scheduling. Under these assumptions, the dynamic scheduling problem is an optimal control problem which can be set as a Markov Decision Process (MDP). This constitutes our starting point. Traditionally, after setting up the MDP, one tries to obtain structural results on the optimal policy by using induction on the time horizon. We choose to proceed in a different direction instead and use coupling and sample path comparison techniques to obtain a partial but exact characterization of the optimal policy.

For a single class make-to-stock queue, the issues are considerably simpler since resource partition is not a concern. In that case, the issue is to determine when to produce and when to idle. It turns out that a base-stock type policy is optimal for this case, the machine should keep on producing whenever the stock are below a certain *safety* level and should idle once that level is attained (Gavish and Graves, 1980).

Kimemia and Gershwin (1983) pioneered the approach of modeling production control problems as optimal control problems. This led to the characterization of the optimal policy for a single class continuous flow model by Bielecki and Kumar (1988). The continuous flow model is closely related to the discrete part type model considered here and involves a constant demand rate and constant production rate but models the randomness through an unreliable machine which can be up or down. For this model, Bielecki and Kumar have shown that the optimal policy is a hedging point policy, where the production facility works at full rate when it is up and the inventory level is below the hedging point (which determines a safety stock level) and at a rate equal to the demand rate when the inventory level reaches the hedging point.

The multi-class make-to-stock queueing problem was first considered Zheng and

Zipkin (1990). Zheng and Zipkin show that in the case of two symmetric products, the performance of a policy that always serves the longest queue is always better than the performance of a FCFS policy. These results were later generalized to multiple products by Zipkin (1995). Wein (1992) proposed a Brownian approximation for the multi-class make-to-stock queueing control problem. The solution of the approximating stochastic control problem provides interesting insights into the structure of the optimal policy, suggesting particularly the optimality of a hedging point policy and a static priority rule when all products are backlogged.

Ha (1997) provides the theoretical justification of some of the ideas suggested by the approximating model of Wein (1992). By considering the infinite horizon discounted cost model and using dynamic programming he proves that a static priority rule is optimal when all producs are backlogged. He also proves that for two part types requiring identical production times, the optimal policy is a hedging point policy, characterized by two switching curves, one curve determining the on-off region for production and the other curve determining the dynamic priority between the part types.

Ha's results suggest that the optimal policy for the multi-class make-to-stock queueing problem (in further generality than proven) is a hedging point policy combined with monotone priority regions that state which of the part-types to produce. On the other hand, even under the restriction of policies to this particular class, one is left with a challenging problem of jointly optimizing the selection of a hedging point and the priority regions for different part types. Veatch and Wein (1996) and Peña-Perez and Zipkin (1997) study this problem and provide effective heuristics based on approximations driven by intuition.

For the continuous flow two part type problem, on the other hand, Srivatsan and Dallery (1998) have recently provided a partial (but exact) characterization of the optimal hedging point policy. This exact characterization prompts the question as to whether the same properties carry over to the conceptually related but considerably different case of the discrete part make-to-stock queue.

Note that, in the single product case, the optimal control problems for the make-to-stock queue and continuous model of Bielecki and Kumar (1991) are strongly related. In fact, for both cases the optimal policy is of the hedging point type. This connection also holds for the multi-product case when the service times do not depend on the part type. Intuitively, the effect of randomness modeled by machine breakdowns is very similar to the effect of randomness modeled by a class-independent server. In the case where the production times depend on the part type, however, the two systems represent randomness in radically different ways. For this second case, the corresponding continuous flow model seems to be the one with part type dependent breakdown rates which has not received a lot of attention so far.

In this paper, we show the surprising result that the partial characterization of the optimal policy as provided by Srivatsan and Dallery (1998) for a continuous flow two-part type system extends to the two-class make to stock queue with part-type dependent production times. The extension turns out to be technically quite involved, partially due to the passage from the continuous to the discrete case, but mainly due to the different way the models capture the randomness. The end result, however, is very simple and intuitively appealing. In a certain region of the state space, the monotone switching curve that separates the priority regions of the two products turns out to be a straight line whose position is expressed by a simple equation. This characterization allows us to generalize the results of Ha (1997) on the structure of the optimal switching curve. It also helps to recognize the advantages and disadvantages of the various heuristic strategies proposed by Veatch and Wein (1996) and Peña-Perez and Zipkin (1997). In particular, we formally justify the good performance of some of these policies developed through intuitive approaches.

We give a formal definition of the problem and the model in section 2. In section 3, we present some properties of the class of policies that we study in this paper, namely, the hedging point class of policies. These properties enable us to obtain the main result on the characterization of the optimal hedging point policies presented in Section 4. In section 5, we give numerical examples as well as discussing and justifying the relative performance of some of the heuristic policies proposed earlier. Finally, our conclusions and suggestions for future research are presented in Section 6.

2 The Optimal Control Problem

2.1 The Model and the Dynamic Scheduling Problem

Consider a production system with a single, flexible, machine that produces two part types (type 1 and type 2), in a make-to-stock mode. Each finished item is placed in its respective inventory. Demands that cannot be met from their respective on-hand inventories are backordered. It is assumed that raw parts are always available in front of the machine. The arrivals of demands to the system occur according to independent Poisson processes with rates λ_i, $i = 1, 2$. The production times of product i are independent and exponentially distributed with rates μ_i.

At any time, one can choose whether to produce part type 1, 2 or to idle the machine. A preemptive discipline is further assumed: the production of a part can be interrupted and resumed. A control policy states the action to take at any time. Since the system is memoryless, for the control of the system we can consider only Markov policies, which only depend on the current state. Let $X_i(t)$ denote the inventory level at time t. We call $X_i(t)$ the surplus (or backlog if demands are backordered) of Part type i. $\boldsymbol{X}(t) = (X_1(t), X_2(t))$ is then the state of the system. Let C_a be the control associated with a Markov policy a. We have:

$$C_a(t) = C_a(\boldsymbol{X}(t)) = \begin{cases} 0 & \text{when the action is to idle} \\ 1 & \text{when the action is to produce type 1} \\ 2 & \text{when the action is to produce type 2} \end{cases}$$

The issue we are interested in is how to control this system to minimize a cost function. We consider a unit holding cost h_i and a unit backorder cost b_i per unit of time for part type i. Here h_i and b_i are non-negative scalars. In the state $(X_1(t), X_2(t))$, the system incurs an instantaneous cost of

$$c(\boldsymbol{X}(t)) = \sum_{i=1}^{2} c_i(X_i(t))$$

where the individual part type costs c_i are

$$c_i(X_i(t)) = \begin{cases} h_i X_i(t) & X_i(t) \geq 0 \\ -b_i X_i(t) & X_i(t) \leq 0 \end{cases}$$

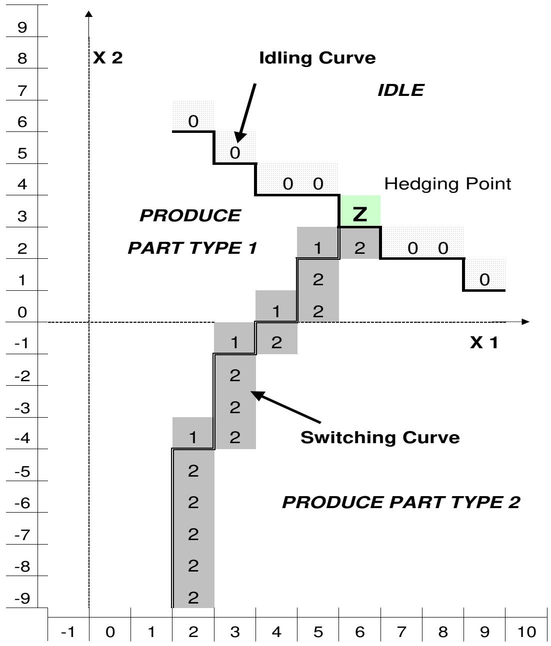

Figure 1: Hedging Point Policy

Under a given policy, say a, the system will generate a long run average cost due to surpluses and backlogs. The objective is then to find the policy which minimizes this long run average cost. This problem can be expressed as an infinite horizon, average cost control problem shown below:

$$\min_{a} \quad \limsup_{t\to\infty} \frac{1}{t} E_x^a[\int_0^t c(\boldsymbol{X}(t))dt] \quad \text{where} \quad x = \boldsymbol{X}(0). \tag{1}$$

2.2 Optimal Control and Hedging Point Policies

To solve the optimal control problem (1), a classical approach is to derive the dynamic programming optimality equations. Following Veatch and Wein (96), with g^* the optimal average cost rate, $V(x)$ the relative value function, we have

$$V(x) + \frac{g^*}{\Lambda} = \frac{c(x) + \lambda_1 V(x_1 - 1) + \lambda_2 V(x_2 - 1) + \mu V(x) + \min(0, \mu_1 \Delta_1 V(x), \mu_2 \Delta_2 V(x))}{\Lambda}, \tag{2}$$

where

$\Delta_1 V(x) = V(x_1 + 1, x_2) - V(x), \Delta_2 V(x) = V(x_1, x_2 + 1) - V(x),$

$\mu = max(\mu_1, \mu_2), \Lambda = \lambda_1 + \lambda_2 + \mu.$

The optimality equation is useful in determining certain structural properties of the optimal policy and also provides the basis for algorithms to compute it numerically. Ha (1996) has exploited these equations to characterize the monotone structure of the optimal policy with discounting. On the other hand, this approach has not given an exact general characterization of the optimal policy until now. Although a formal proof does not exist in full generality, we conjecture that the optimal policy belongs to a specific class : the hedging point (base stock) policies (see Figure 1).

Following Kimemia and Gershwin (1983) and Veatch and Wein (1994), we define a *hedging point policy* for the manufacturing system in Section 2.1 as follows:

Definition 1 *A* hedging point policy *a is a Markov control policy that has the following properties:*

1. *The policy divides the two dimensional surplus space into three distinct types of regions, say Regions* $0, 1$ *and* 2 *such that* $C_a = 0$ *in Region Type* 0*;* $C_a = 1$ *in Region Type* 1*;* $C_a = 2$ *in Region Type* 2*.*

2. *The idling curve defined as the set of points of Region* 0 *on the boundary between the Region* 0 *and the others, is a positive curve, decreasing in* x_1 *or* x_2*. (The curve can be written as a function of* x_1 *in the values of* x_2*, or of* x_2 *in the values of* x_1*).*

3. *The switching curve which delimits Regions* 1 *and* 2*, is the following increasing in* x_1 *or* x_2 *curve defined by* $\{\boldsymbol{x} : (x_1 - 1, x_2) \in$ *Region* 1 *and* $(x_1, x_2 - 1) \in$ *Region* $2\}$*. (The curve can be written as a function of* x_1 *in the values of* x_2*, or of* x_2 *in the values of* x_1*).*

4. *These two curves intersect at a single point* $z = (z_1, z_2)$*, which is called the* hedging point.

Remark: The parts 2 and 3 of the definition imply that the hedging point z is positive.

When a hedging point policy is applied, if the initial point (x_1, x_2) of a surplus trajectory is such that $x_1 \leq z_1$ and $x_2 \leq z_2$, the machine will be used at full capacity until the inventories levels are equal to z. Thus, the hedging point represents a surplus level based on tradeoffs between expected inventory and backlog costs. Furthermore, if at a time instant the trajectory reaches the switching curve at a certain point y, it will move along the curve until an arrival of a new demand causes it to move away from the curve. Then the scenario is repeated until z is reached. This implies that all the points of the switching curve between y and the hedging point z belong to the trajectory.

The switching curve, with the hedging point, fully characterize a hedging point policy for initial points verifying the above condition. In particular, if a surplus trajectory originates at the hedging point, its future behavior can be completely characterized by these parameters.

Furthermore, if $x_1 > z_1$ or $x_2 > z_2$ then after a transient behavior which does not affect the long average cost, a trajectory under a hedging point policy will reach a point such that $x_1 \leq z_1$ and $x_2 \leq z_2$. The description above then holds true.

In the rest of the paper we concentrate on hedging point policies, with starting points in the region $x_1 \leq z_1$ and $x_2 \leq z_2$. Although a rigorous proof under general assumptions is not yet available, there are strong evidences that the optimal policy is a hedging point policy.

Previous work on this problem supports this conjecture. Firstly, Ha (1996) has shown that the optimal policy in the discounted case is a hedging point policy in the case of $\mu_1 = \mu_2$. Hedging point policies are "plausible" according to Peña-Perez and Zipkin (1997). This issue is also discussed from the monotonicity point of view in Veatch and Wein (1996). Finally, our numerical experiments verify this point.

2.3 An Equivalent Model

We consider now a model which only differs from the original model in the way the machine produces the parts. This new model, whose behavior will be shown to be equivalent to that of the original model will be useful for deriving some of our results.

Definition 2 *The Equivalent model (EQ) is a model similar to the original model where:*

1. *the machine performs service activities whose durations do not depend on the type of the product and are exponentialy distributed with rates* $\mu = \mu_1 + \mu_2$*,*

2. *at the end of a service time the "work" done during the service activity is either allocated to one of the two parts or not used, according to a control policy a. Specifically, we have*

$$C_a = \begin{cases} 0 & \textit{idle: the work is not used} \\ 1 & \textit{produce type 1: the work is allocated to type 1} \\ 2 & \textit{produce type 2: the work is allocated to type 2} \end{cases}$$

3. *When $C_a(t) \neq 0$, the allocation of the "work" to a given part type may result in the "instantaneaous production" of a part of this type, and the outcome is probabilistic. Specifically if the work is allocated to part type i, then with probability $p_i = \frac{\mu_i}{\mu}$, a part is instantaneously delivered to the corresponding output buffer, while with probability $1 - p_i$ nothing happens, i.e the "work" is lost.*

Remark: In the EQ model we talk about "service times" instead of "production times". Note that not all service times correspond to actual production times.

An intuitive interpretation of this model is that the machine is not perfectly reliable: it can produce parts which do not satisfy some quality criteria. For instance with a probability $1-p_i$, a part of type i is "bad" and is rejected when it is produced.

Since the arrival processes and the service times are independent of the policy, this model facilitates sample path comparisons, as we will see later.

Proposition 1 *Under the same control policy, the behavior of the EQ model is probabilistically equivalent to that of the original model.*

Proof: Consider first the machine and the both production processes. Suppose then that the common policy a states to produce part i. Let U_i be the discrete random variable such that:

$$U_i = \begin{cases} 1 & \text{when the work allocated to part type } i \text{ generates a real part} \\ 0 & \text{when the work does not generate a real part (the work is lost)} \end{cases}$$

with the probabilities:

$$P(\{U_i = 1\}) = p_i \text{ and } P(\{U_i = 0\}) = 1 - p_i.$$

U_i corresponds to the probabilistic outcome mentioned in the third part of the EQ model definition. Let T and T_i, be the exponentially distributed service times with

rates μ and μ_i respectively. Consider the following events:
A={in the original model, the service time of part type i is less than dt}
B={in the EQ model, the service time of part type i is less than dt}.

We have:

$$P(A) = P(T_i < dt) = \mu_i dt + o(dt) = p_i \mu dt + o(dt) = P\left(\{U_i = 1\} \cap \{T < dt\}\right) = P(B).$$

When a states not to produce, nothing happens in the original as well as in the EQ model. Thus under the same policy, the two models have a stochastically identical behavior for their service time process.

Consider now the arrival processes. When a demand occurs, the effect on the inventory level is exactly the same for both models. In the original model, the attribution of the part type to the machine is determined by the preemptive discipline. It is possible to switch part types at the time instant of the arrival, even when the machine is working. In the EQ model, since we choose to attribute the part after its completion, the attribution has not yet been made. The effects of the arrival processes are thus exactly the same for both models. $\square$

Note that the EQ Model corresponds to a uniformization of the service processes in the Original Model. The proof of the above property is given here for clarity and to introduce some useful notation to be used in the sequel.

3 Some Properties of Hedging Point Policies

3.1 Sample Path Properties

This subsection presents some properties of hedging point policies. These properties will enable us to study the stability conditions. Moreover, they will also constitute the basis of the proof for the characterization of optimal hedging point policies. These results are based, as others in this paper, on sample path comparisons. We study different trajectories by coupling them, that is by considering a common realization of the random variables which generate them. In particular, we typically consider the trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ generated by the policies a and b under the same realizations of the arrival processes and the service times.

Following this approach, the next lemma states that if the same number of type-1 parts and the same number of type-2 parts have been completed for two coupled trajectories, then the completion instances of the batches must be identical. Thus, under certain conditions the lemma provides us the positions of coupled trajectories at the same instance which will prove to be very critical for the sample path comparisons in the sequel.

Lemma 1 *Consider two hedging point policies a, b and two coupled trajectories $\boldsymbol{X}^a$, $\boldsymbol{X}^b$ generated by these policies. If two time instants T_a and T_b are such that:*

1. $X_1^a(T_a) - X_1^a(0) = X_1^b(T_b) - X_1^b(0)$, *and* $X_2^a(T_a) - X_2^a(0) = X_2^b(T_b) - X_2^b(0)$
2. *For all t in $[0, T_a]$ (respectively $[0, T_b]$) with $i = 1$ or $i = 2$,*

$$X_i^a(t) \leq X_i^a(T_a) (\textit{respectively } X_i^b(t) \leq X_i^b(T_b))$$

   and T_a (resp. T_b) is the first time that $\boldsymbol{X}^a$ (resp. $\boldsymbol{X}^b$) reaches $\boldsymbol{X}^a(T_a)$ (resp $\boldsymbol{X}^b(T_b)$).

3. *For all t in $[0, max(T_a, T_b)]$, the machine works at full capacity under both policies,*

then $T_a = T_b$ on the coupled path.

Proof: Consider the EQ model. Define T and U_i as in the proof of Theorem 1. We couple the service time T of the machine and the random variables U_i for the policies a and b.

Consider a sequence of realizations of the exponentially distributed service times:

$$t^1, t^2, \ldots, t^n, \ldots$$

and sequences of realizations of the discrete random variables U_1:

$$u_1^1, u_1^2, \ldots u_1^n, \ldots \text{ where } u_n^1 = 0 \text{ or } 1$$

and U_2:

$$u_2^1, u_2^2, \ldots u_2^n, \ldots \text{ where } u_n^2 = 0 \text{ or } 1$$

These realizations are common for both policies. Let T^n be the time instant of the nth event. Since the machine works at full capacity, $T^n = \sum_{k=1}^{n} t^k$, and T^n is the same for both policies. We denote by $n(t)$ the number of service completions at time t. At instant T^n, the nature of the new event is then given by the next element of the sequence corresponding to u_i (where $i = 1$ or $i = 2$ depending on the choice stated by the policy). Let $n_i^a(t)$ be the number of elements of the sequence corresponding to u_i, which policy a has chosen before the time instant t. Similarly, we can define n_i^b for policy b. From the definition of n_i^a, and since the machine works at full capacity, we have for all t in $[0, max(T_a, T_b)]$,

$$n(t) = n_1^a(t) + n_2^a(t) = n_1^b(t) + n_2^b(t). \tag{3}$$

Let $d_i(t)$ be the number of demands for part Type i, which have occurred before the instant t. By coupling, these arrivals modify the inventory level at the same instant for both policies. A demand may then change the choice stated by the policy, but for all t in $[0, max(T_a, T_b)]$, $d_i(t)$ stays the same for a and b. Each time n_i^a is equal to 1, the corresponding event is a production of a part. Thus,

$$X_i^a(t) - X_i^a(0) = \sum_{k=1}^{n_i^a(t)} u_i^k - d_i(t). \tag{4}$$

Suppose now that $T_a < T_b$. From the second condition of the lemma, it follows that for $i = 1$ or $i = 2$,

$$X_i^b(T_b) - X_i^b(0) \geq X_i^b(T_a) - X_i^b(0), \tag{5}$$

and one of the inequalities is strict. Without loss of generality suppose that,

$$X_1^b(T_b) - X_1^b(0) > X_1^b(T_a) - X_1^b(0), \tag{6}$$

From the first condition of the lemma, and from (5) and (6), we obtain,

$$X_1^a(T_a) - X_1^a(0) \quad > \quad X_1^b(T_a) - X_1^b(0) \tag{7}$$

$$X_2^a(T_a) - X_2^a(0) \quad \geq \quad X_2^b(T_a) - X_2^b(0) \tag{8}$$

By combining (7) and (8) with (4) we obtain the following results :

$$n_1^a(T_a) > n_1^b(T_a) \text{ and } n_2^a(T_a) \geq n_2^b(T_a)$$

which from (3) are impossible. Using similar arguments when $T_a < T_b$, we have $T_a = T_b$. $\square$

The previous lemma gives general conditions for two trajectories to complete the same amount of work within the same time. This result will be adapted to the class of policies defined by hedging points and switching curves in the following two corollaries.

Corollary 1 *Consider two hedging point policies a, b and two coupled trajectories $\boldsymbol{X}^a$, $\boldsymbol{X}^b$ generated by these policies such that,*

1. *the switching curves of the policies have a common point $\hat{\boldsymbol{z}}$,*

2. *$\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ start at the same initial point, such that $X_1(0) \leq \hat{z}_1$ and $X_2(0) \leq \hat{z}_2$,*

then, the trajectories reach $\hat{z}$ at the same time instant.

Proof: As explained in Section 2.2, once the trajectory reaches the switching curve, it moves along it. From condition 2 above, both trajectories will reach their respective switching curve below $\hat{\boldsymbol{z}}$. Thus both trajectories pass through this point.
Both policies also work at full capacity and Lemma 1 can be applied with $\boldsymbol{X}^a(0) = \boldsymbol{X}^b(0)$ and $\boldsymbol{X}^a(T_a) = \boldsymbol{X}^b(T_b) = \hat{\boldsymbol{z}}$. $\square$

Corollary 2 *Consider two hedging point policies and two coupled trajectories beginning at their respective hedging points. The inventory levels leave and reach their respective hedging points at time instants which are the same for both policies.*

Proof: At the time instant of the first arrival, both trajectories will decrease by 1 in the respective inventory level. Then, they verify the lemma 1. Thus they will reach their hedging point at the same instant, and the scenario will be repeated. $\square$

3.2 Stability of Hedging Point Policies

In this subsection we analyse the stability of the hedging point policies. We define Policy a with the hedging point z, to be stable if

$$\lim_{t\to\infty} \frac{1}{t} E_z^a[\boldsymbol{X}(t)^2] = 0 \quad \text{for} \quad i = 1, i = 2.$$

Corollary 2 enables us to relate the stability of hedging point policies to the stability of static priority policies as presented in the following lemma.

Lemma 2 *A hedging point policy a is stable if*

$$\frac{\lambda_1}{\mu_1} + \frac{\lambda_2}{\mu_2} < 1$$

Proof Let z^a be the hedging point of the policy a. Let Policy b, be the policy with the same hedging point z^a, and the switching curve given by the straight-line $x_1 = z^a$. From Corollary 2, Policy a is stable if and only if Policy b is stable. Policy b actually defines a static priority policy with part type 1 having the high priority. Thus, following Gross and Harris (1985), Policy b is stable if $\lambda_1/\mu_1 + \lambda_2/\mu_2 < 1$ giving us the result. □

Remark: Note that the stability condition of hedging point policies does not depend on the hedging point or the switching curve but only on the quantity $\lambda_1/\mu_1 + \lambda_2/\mu_2 < 1$, which can be interpreted as the utilization rate of the machine.

4 Partial Characterization of the Optimal Hedging Point Policy

Consider the two-part-type system introduced in Section 2.1. Without loss of generality, let the two part types be numbered such that

$$b_1\mu_1 \geq b_2\mu_2 \tag{9}$$

We derive in this section, a structural result for the optimal hedging point policy of the two-part-type system introduced in Section 2.1. The main idea is to relate the optimal control problem of the two-part-type system to a single-part-type system problem. Informally, we will exploit the fact that the instantaneous cost function can be expressed as:

$$c(x_1, x_2) = c^m(x_1) - f(x_1, x_2) \text{ for } x_2 \leq 0$$

where c^m is the part of the cost that only depends on x_1, and f is a function of x_1 and x_2 that captures the remaining part of the cost. c^m is an instantaneous cost function of the form:

$$c^m(x_1) = \begin{cases} h^m x_1 & x_1 > 0 \\ -b^m x_1 & x_1 \leq 0 \end{cases}$$

Intuitively, the function f should be proportional to W, the total amount of work (in units of time) embodied in the system.

Definition 3 *Let the aggregate workload $W(t)$ be defined as:*

$$W(t) = \frac{X_1(t)}{\mu_1} + \frac{X_2(t)}{\mu_2}$$

We denote by $W^a(t)$, the aggregate workload under policy a. We also use the following notations $W(\boldsymbol{X}) = W(\boldsymbol{X}(t)) = W(t)$. Under certain conditions, the expectation of $W(t)$ does not depend on the policy, or differs only by a constant. Thus, the difference in average costs of two given policies can be resumed by the difference in $c^m(x_1)$, which is the cost of a single part-type system. We are then able to give an analytical expression for the switching curve when $x_2 < 0$.

Theorem 1 formalizes this characterization of the optimal policy. The following lemmas give the properties of the expected value of the aggregate workload mentioned above.

Lemma 3 *Consider two trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ generated by two hedging point policies a and b, and an interval $[0, T]$ such that for all t in $[0, T]$ the machine works at full capacity,*
then for all t in $[0, T]$ we have,

$$\Delta_b^a E[W(t)] = E^a[W(t)] - E^b[W(t)] = W^a(0) - W^b(0)$$

Proof: Consider the EQ model. Let T^n be the time instant random variable of the nth event, which can be a demand arrival or a service completion (recall that in the EQ model, a service completion does not necessarily correspond to a production completion). For a policy π we denote by $H_n = (X_0, C_\pi(0) \ldots X_n, C_\pi(T_n))$ the history of the processes until the time of the nth event, where C_π is the control associated ($C_\pi = 1$ or 2). We also denote $\hat{E}^\pi[W(t)]$ to be equal to $E^\pi[W(t)|H_n]$, with $T^{n-1} < t \leq T^n$. We have,

$$E^\pi[W(t)] = E[\hat{E}^\pi[W(t)]]. \tag{10}$$

Without loss of generality, we can take t such that $t = T^n$.

Suppose then that the instant T^n corresponds to a service completion. It follows for policy a when $C_a(T_n) = i$ that,

$$\hat{E}^a[W(T^n)] = \hat{E}^a[W(T^{n-1}) + \frac{U_i}{\mu_i}] \tag{11}$$

where U_i is the discrete random variable of the EQ model. From the definition of U_i (see the proof Proposition 1) we obtain that $E[U_i]/\mu_i = 1/\mu$. Note that this value does not depend on the part type. Thus, from (11) we have,

$$\hat{E}^a[W(T^n)] = \hat{E}^a[W(T^{n-1})] + \frac{1}{\mu}. \tag{12}$$

If T^n corresponds to an arrival of type i, then

$$\hat{E}^a[W(T^n)] = \hat{E}^a[W(T^{n-1})] - \frac{1}{\mu_i} \tag{13}$$

Similary for policy b:

$$\hat{E}^b[W(T^n)] = \hat{E}^b[W(T^{n-1})] - \frac{1}{\mu} \text{ or } \hat{E}^b[W(T^n)] = \hat{E}^b[W(t^{n-1})] - \frac{1}{\mu_i} \tag{14}$$

Consider now realizations t_i of T^i for $i \in \{0, \ldots, n\}$. As it is done in the proof of lemma 1, we couple the service times of the machine as well as the instants of demand arrivals. Thus the t_i instants are the same under policy a and b. From (12), (13), (14), it follows then that:

$$\Delta_b^a \hat{E}[W(T^n)] = \Delta_b^a \hat{E}[W(T^{n-1})]$$

giving us the desired result using (10). □

Remarks: In the case where $\mu_1 = \mu_2$, we actually have $W^a(t) - W^b(t) = W^a(0) - W^b(0)$ on the coupled path.

By combining the two lemmas, we obtain a property which holds for any time instant, even when the machine is not working.

Lemma 4 *Consider two hedging point policies begining at their respective hedging points z^a, z^b. Then, for all t, $\Delta_b^a E[W(t)] = \frac{z_1^a}{\mu_1} + \frac{z_2^a}{\mu_2} - \frac{z_1^b}{\mu_1} - \frac{z_2^b}{\mu_2}$.*

Proof: When the policies are hedging point policies, condition 2 of Lemma 3 holds until one of the two trajectories reaches its hedging point. Using corollary 2, the property holds then for all time instants t. □

Based on lemma 1, 3, and 4 we can now formulate the following theorem which gives an analytical expression for the switching curve in a certain region of the space.

Theorem 1 *Consider a two-part-type system where (9) holds as a strict inequality. When $x_2 < 0$, the switching curve of the optimal hedging point policy for this system is the straight-line defined by:*

$$x_1 = z_1^m = \lfloor \frac{\ln\left(\frac{h_1 + b_2 \frac{\mu_2}{\mu_1}}{h_1 + b_1}\right)}{\ln \frac{\lambda_1}{\mu_1}} \rfloor$$

and the optimal hedging point policy a is of the form:

$$C_a(\boldsymbol{x}) = \begin{cases} 1 & \text{if } x_1 < z_1^m, x_2 < 0 \\ 2 & \text{if } x_1 \geq z_1^m, x_2 < 0 \end{cases}$$

Proof: A proof of this theorem can be found in the appendix.
Remark: Note that when $x_2 < 0$, the switching curve depends neither on the arrival process of type 2 demands, nor on the holding cost of part Type 2.

It is interesting to note that the value of z_1^m as calculated above, could be zero for certain a range of parameters. A direct calculation gives then the following property:

$$z_1^m = 0 \iff h_1\mu_1 - h_1\lambda_1 > b_1\lambda_1 - b_2\mu_2.$$

This property leads to an intuitive interpretation of Theorem 1. Consider the case where $z_1^m > 0$ and an inventory level with $z_1^m > x_1 \geq 0$ and $x_2 < 0$. From Theorem (1), the optimal policy states to produce one unit of product 1 which increases the expected cost at rate $h_1\mu_1 - h_1\lambda_1 - h_2\lambda_2$. While, if the policy had stated to produce one unit of product 2, the expected cost would have increased at rate $-b_2\mu_2 + b_1\lambda_1 - h_2\lambda_2$. Thus, the optimal policy chooses the part to minimize this cost-change rate. This extends the result and the interpretation given by Ha (1997) to justify the optimality of the "$b\mu$" rule.

5 Numerical Results and Heuristic Policies

5.1 Numerical Results

To numerically compute optimal policies, we have used the value iteration algorithm to solve the optimality equation (2). To overcome the problem of infinite state space,

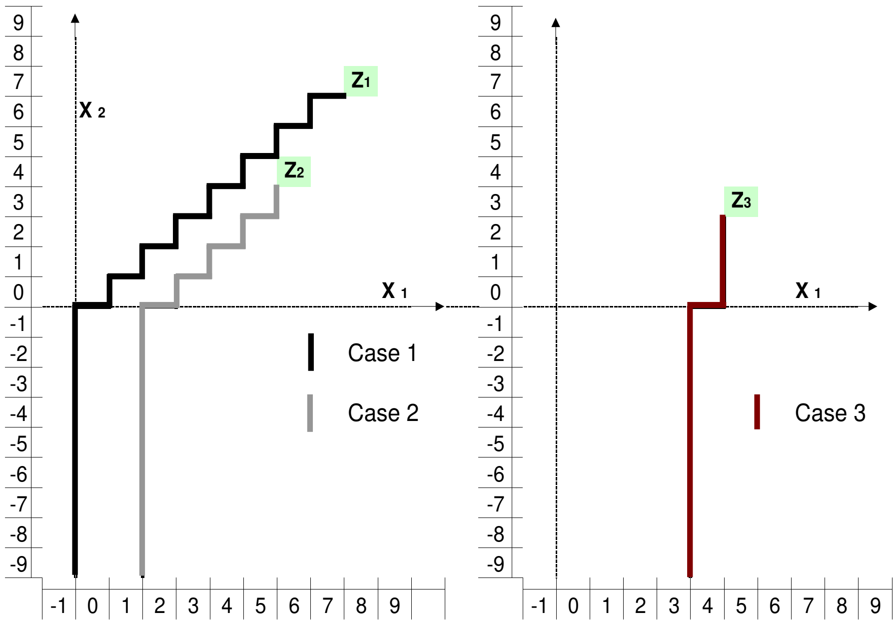

Figure 2: Numerical results

we truncated the state space and experimented with increasingly large state spaces until the increase did not cause any change.

We perform a computional experiment to provide some insights on the behavior of the optimal policy when $x_2 < 0$ varying the parameters of the system. The data of the different problems we have studied are displayed in Table I. In the two first cases, we consider systems which only differ by their backlog costs. The ratio b_1/b_2 increases from the first case to the second one. In the last case, the system is also asymmetric in the arrival and service time process, such that $\rho_1 = 2\rho_2$.

In Figure 2 which displays the optimal policy for the examples in Table I, the straight line appears clearly. The figures also verify that the experimental value of this line is equal to the theorical value given by Theorem 1, which is reported as Z^m in Table I.

Table I: Data of the tested cases

Case	λ_1	λ_2	μ_1	μ_2	h_1	h_2	b_1	b_2	Z^m
1	0.4	0.4	1	1	1	1	50	25	0
2	0.4	0.4	1	1	1	1	50	5	2
3	0.6	0.1	1.2	0.4	2	1	60	6	4

Let us decribe the behavior of this line (in all this discussion we assume that $b_1\mu_1 \geq b_2\mu_2$). Consider first that $\mu_1 = \mu_2 = 1$ as in cases 1 and 2. The position of the switching curve when $x_2 < 0$ depends then on b_1, b_2 and h_1. More precisely, it depends on the relative values b_1/h_1, b_2/h_1. When h_1 tends to infinity, the line is at 0. Note however that in practice h_1 is usually less than b_1. Suppose then for simplicity that $h_1 = 1$, as in the two first cases. Z^m increases in the ratio b_1/b_2. Thus the more the two parts are assymetric in their backlog costs, the larger the stock of part type 1 when $x_2 < 0$ must be. This can be seen in Figure 2. When $\mu_1 \neq \mu_2$, the behavior of the line is the same by taking the cost rate $b_i\mu_i$ instead of b_i with $i = 1$ or $i = 2$, and $h_1\mu_1$ instead of h_1.

In addition to the asymetry of the backlog cost rates, the utilisation ratio of the first part type has a crucial impact in the position of the switching curve. Indeed, Z^m increases with ρ_1. Particulary, in heavy traffic, the position of the line tends to infinity. In the third case, we have a straight line at $Z^m = 4$ when ρ_1 equals only 0.5. Here Z^m is very close to the hedging point of part Type 1 (which from Figure 2 is equal to five).

Note that for a certain set of parameters, the optimal policy can be a static priority policy. In that case, Z^m is equal to the coordinate of the hedging point corresponding to the first part type.

Thus, the position of the optimal switching curve when $x_2 < 0$ seems to capture at once the asymmetry of the backlog cost rate, and the utilization in isolation of the part type with the larger of these costs.

5.2 Myopic Allocation

In Section 2, we have seen that no exact solution has been found for the Dynamic Scheduling Problem (1). Consequently, efforts have been devoted to explore heuristic approaches.

For instance, a simple heuristic policy is a static priority policy with a hedging point z, which switching curve is defined by the straight-line $x_1 = z_1$. Computations can then give an approximation of the optimal hedging point z for this class of policies. These heuristics have been studied by Wein (1992) and have been called by Peña-Perez and Zipkin (1997) the static-priority(r) Policy (with $r_i = z_i/\mu_i$). Using Theorem 1, we can see that this kind of policy would perform efficiently in cases where the hedging point of part Type 1 is close to the position of the straight-line as it can be seen in the third case of our numerical results. However, in the more general case, the static priority will not perform well.

Thus, other more sophisticated, yet easily computable heuristics have been explored. In particular, Peña-Perez and Zipkin (1997) have developed heuristics (the "Myopic Allocation") which perform substantially better than the static-priority(r) Policy. Veatch and Wein (1996) have also studied these heuristics coupled with a Brownian approximation developed by Wein (1992). They show that these myopic allocation policies give very good results when applied to approximate the optimal switching curve. However, the myopic allocation is based on intuitive but informal arguments. After its short presentation, we give a partial justification of the myopic allocation using Theorem 1.

The main idea of the myopic allocation, is to look-ahead a service time of part type i, say S_i. The policy allocates then the production capacity to the part type which increases the expected instantaneous cost at the smaller rate. Let $D(S_i)$ be the number of demands of part type i in the interval $[0, S_i]$. If part type i is produced and if the current inventory level is x_i, $g(x_i) = E[c(x_i+1-D(S_i))]$ is then the expected instantaneous cost after the completion of the service time. Thus, $\mu_i \Delta g(x_i) = g(x_i+1) - g(x_i)$ is the rate at which serving this class increases the instantaneous expected cost. The myopic allocation chooses then the part type with the smaller $\mu_i \Delta g(x_i)$. Peña and Zipkin have derived an analytical expression for this index $\mu_i \Delta g(x_i)$. However, they suggest that, taking instead of S_i the sojourn time T_i in a system (where $j \neq i$ are omitted), the heuristic policy will perform better. Thus in that case, $D(T_i)$ replaces $D(S_i)$ in the computations above, where T_i is exponentially distributed with rate $(1 - \rho_i)\mu_i$. The intuitive rational they give is that, by choosing product type i, the production requirement is to increase x_i by one taking into account the demands that arrive in the mean time. The machine will then not only produce one part type i, but will also respond to the new demands until the inventory level reaches $x_i + 1$. This "replenishment" time has actually a mean proportional to $(1-\rho_i)^{-1}$, as has $E(T_i)$. The improvement brought by this modification was verified by their numeric examples. Indeed, they remark that the myopic policy with T_i performs better than the myopic policy with S_i, which is better than the static priority rule. But they do not give any formal justification of these conclusions.

The following Lemma partially proves that their intuitions were well-founded.

Lemma 5 *The myopic allocation policy applied with the sojourn time T_i is an optimal policy when $x_2 < 0$.*

Proof: Let (x_1, x_2) be an inventory level. Without loss of generality we suppose that $b_1\mu_1 > b_2\mu_2$. The myopic allocation policy chooses the part type with the smaller $\mu_i \Delta g(x_i)$ ($i = 1$ or $i = 2$). Following Peña and Zipkin (1997), we have with the sojourn time T_i,

$$\mu_i \Delta g(x_i) = \begin{cases} -\mu_i b_i & \text{for} \quad x_i < 0 \\ -\mu_i b_i + \mu_i (h_i + b_i)(1 - \rho_i^{x_i+1}) & \text{for} \quad x_i \geq 0 \end{cases}$$

When $x_1 < 0$ and $x_2 < 0$, since $-b_1\mu_1 < -b_2\mu_2$, part type 1 is produced. When $x_1 \geq 0$ and $x_2 < 0$, part type 1 is still produced as long as $\mu_1 \Delta g(x_1) < \mu_2 \Delta g(x_2)$ which can be written as,

$$-\mu_1 b_1 + \mu_1 (h_1 + b_1)(1 - \rho_1^{x_1+1}) \quad < \quad -\mu_2 b_2.$$

A straightfoward computation shows that the myopic allocation policy states to produce part type 1 when $x_2 < 0$ if and only if,

$$x_1 \quad < \quad \lfloor \frac{\ln\left(\frac{h_1 + b_2 \frac{\mu_2}{\mu_1}}{h_1 + b_1}\right)}{\ln \frac{\lambda_1}{\mu_1}} \rfloor.$$

Using Theorem 1, the myopic allocation policy with T_i is thus optimal when $x_2 < 0$. □

This Lemma explains the good performance of the myopic allocation computed with the sojourn time. A direct calculation will show that, for the myopic allocation with the service time, the switching curve is also a straight-line when $x_2 < 0$ with

$$x_1 = \lfloor \frac{\ln\left(\frac{h_1 + b_2 \frac{\mu_2}{\mu_1}}{h_1 + b_1}\right)}{\ln \frac{\lambda_1}{\mu_1 + \lambda_1}} \rfloor. \tag{15}$$

Thus in light traffic conditions, the myopic policy with S_i is close to the optimal policy when $x_2 < 0$. However in heavy traffic condition for part type 1, the straight-line is at $x_1 = 0$, while this line tends to infinity for the optimal policy. This explains in part that the myopic policy with the service time can perform poorly in contrast with the one that uses the sojourn time.

6 Conclusion

Using sample path comparisons for hedging point policies, we have partially characterized the switching curve that determines the production priorities for the two-class make-to-stock queue. Our results suggest that in the case where both products are backlogged, it is optimal to produce the most expensive item in terms of the backorder cost (the product with the higher $b\mu$) until its stock reaches a predetermined (non-negative) level before switching to save the less expensive product from backlog. In addition, it is shown that this safety stock level does not depend on the level of backlogs of the less expensive products and can, in certain cases, be significantly high depending on the cost and traffic parameters.

Similar results have been shown for an analogous continuous flow model. On the one hand, it may be considered somewhat surprising that the optimal policy should have the identical structure for the make-to-stock queue as for the two part type continuous model with an unreliable machine. On the other hand, it is relieving that the optimal policy has the same structure for two models that represent randomness in very different ways indicating the robustness of the structure. This strongly suggests, for instance, that for the continuous model with part type dependent breakdown rates the structure should be retained.

Our results contribute to the understanding of the control problem of the single stage multi-product system. These results could also be useful for the multi-product multi-stage system which constitutes a major challenge both from theoretical and practical perspectives. Future research will focus on some of the issues in multiple stage production.

References

[1] T. Bielecki and P.R Kumar (1988), Optimality of Zero Inventory Policies for Unreliable Manufacturing Systems. *Operations Research*, Vol. 36 pp 532-546.

[2] J.A. Buzacott and J. G. Shanthikumar, *Stochastic Models of Manufacturing Systems*, Prentice Hall, 1993.

[3] B. Gavish and S. Graves (1980), A One-Product Production/Inventory Problem under Continuous Review Policy, *Operations Research* 28, pp 1228-1236.

[4] D. Gross and C.M. Harris, *Fundementals of Queueing Theory*, John Wiley and Sons, 1985.

[5] A. Ha (1997), Optimal Dynamic Scheduling Policy for a Make-to-Stock Production System, *Operations Research* 45, pp 42-53.

[6] J. Kimemia and S. B. Gershwin (1983), An Algorithm for Computer Control of a Flexible Manufacturing System, *IIE Transactions*, pp 353-362.

[7] A. Peña Perez and P. Zipkin (1997), Dynamic Scheduling Rules For a Multiproduct Make-to-Stock Queue, *Operations Research* 45, pp 919-930.

[8] N. Srivatsan and Y. Dallery (1998) Partial Characterizations of Optimal Hedging Point Policies in Unreliable Two-Part-Type Systems, *Operations Research* 46, pp 36-45.

[9] M. Veatch and L.M. Wein (1996), Scheduling a Make-to-Stock Queue: Index Policies and Hedging Points, *Operations Research* 44, pp 634-647.

[10] L.M. Wein (1992), Dynamic Scheduling of a Multiclass Make-to-Stock Queue, *Operations Research* 40, pp 724-735.

[11] Y. Zheng and P. Zipkin (1995), A Queueing Model to Analyze the Value of Centralized Inventory Information, *Operations Research* 38, pp 296-307.

[12] P. Zipkin (1995), Performance Analysis of a Multi-Item Production-Inventory System under Alternative Policies, *Management Science* 41, pp 690-703.

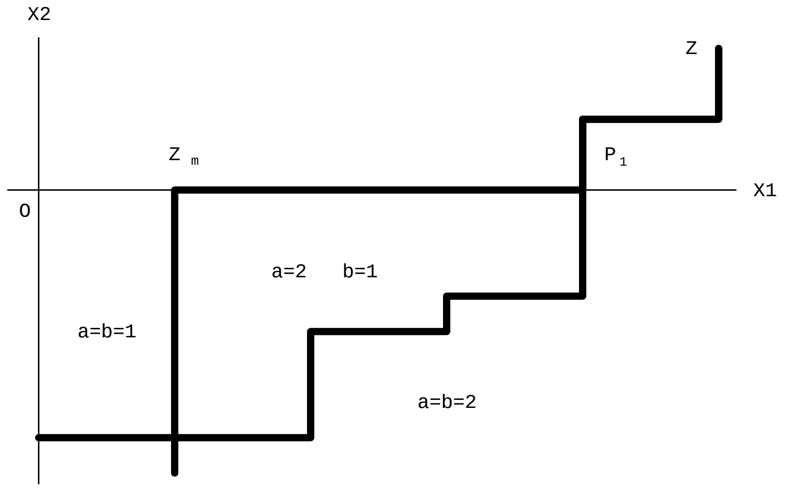

Figure 3: Case 1 $p_1 > z_1^m$

A Appendix

A.1 Proof of Theorem 1

Let Policy b be an optimal hedging point policy with a non-negative hedging point, $\boldsymbol{z} = (z_1, z_2)$, and a switching curve which differs from the straigth-line defined in Theorem 1. This switching curve for Policy b has at least one point on the x_1 axis. Let p_1 be the minimum of the x_1 coordinates of these points.

There are two cases to be considered depending on whether p_1 is greater than z_1^m or not.

Case 1**:** $p_1 > z_1^m$

Let us construct another hedging point policy, Policy a, with the same hedging point z as Policy b, and a switching curve which is: a vertical line through $(z_1^m, 0)$ for $x_2 < 0$; the x_1 axis for $x_2 = 0$, $z_1^m \leq x_1 \leq p_1$; the same as the switching curve of Policy b elsewhere. Policies a and b are illustrated in Figure 3. Consider two trajectories, $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$, that start at the hedging point and evolve under Policies a and b respectively. These trajectories are the same until an arrival of type 2 causes them to enter the region $x_2 < 0$. Let s denote the last time instant just before the two trajectories separate. Since $X_1^a(s) = X_1^b(s) < p_1$, from corollary 1, both trajectories reach the point $(p_1, 0)$ at the same time instant T_1.

By construction, Trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ are identical for $t > T_1$ until an arrival of type 2 again takes them to the region $x_2 < 0$ in which case the above scenario restarts.

Consider now the costs of Policies a and b over the interval $[s, T_1]$.
From Section 2.1, the instantaneous cost function for $x_2 < 0$, is given by:

$$c(\boldsymbol{x}) = \begin{cases} h_1 x_1 - b_2 x_2 & x_1 > 0 \\ -b_1 x_1 - b_2 x_2 & x_1 \leq 0 \end{cases}$$

This can also be expressed as:

$$c(x_1, x_2) = c^m(x_1) - b_2 \mu_2 W(x_1, x_2) \text{ for } x_2 \leq 0$$

where W is the aggregate workload, and c^m is the instantaneous cost function given by:

$$c^m(x_1) = \begin{cases} h^m x_1 = & (h_1 + b_2 \frac{\mu_2}{\mu_1}) x_1 & x_1 > 0 \\ -b^m x_1 = & -(b_1 - b_2 \frac{\mu_2}{\mu_1}) x_1 & x_1 \leq 0 \end{cases}$$

Since hedging point policies are stable, it follows that,

$$\begin{aligned} \Delta_b^a E_x[\int_0^T c(\boldsymbol{X}(t)) dt] &= \int_0^T \Delta_b^a E_x[c(\boldsymbol{X}(t))] dt \\ &= \int_0^T \Delta_b^a E_x[c^m(X_1(t))] dt \; - \; b_2 \mu_2 \int_0^T \Delta_b^a E_x[W(t)] dt \end{aligned}$$

At time s we have $W^a(s) = W^b(s)$, and for all t in $[0, T_1]$, the machine works at full capacity. Thus for t in $[0, T_1]$, $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ verify the conditions of lemma 3 and, we obtain $\Delta_b^a E_x[W(t)] = 0$.

So the difference in expectations of cost between the two trajectories in $[s, T]$ is the same as that for system where the instantaneous cost function in the non-positive x_2 region is given by c^m. It can be noted that this cost function depends only on the value of the Part Type 1 surplus, x_1.

Consider the behavior of X_1 over $[s, t_1]$ when $X_2 < 0$, under Policies a and b respectively. Under Policy a, Part Type 1 behaves as if it were following the policy given by:

$$a_1 = \begin{cases} 1 & x_1^a < z_1^m \\ 0 & x_1^a = z_1^m \end{cases}$$

Thus the surplus trajectory for Part Type 1 under Policy a over the interval $[s, t_1]$ when $X_2^a < 0$ is the same as that for generated by an optimal hedging point policy for the single-part-system where the arrival rate, the production rate and the function of cost are respectively, λ_1, μ_1 and c^m. One can see that the corresponding optimal hedging value is given by z_1^m (see equation (20) of the following appendix). Also, by construction, when $X_2^a(t) = 0$ we have

$$x_1^a(t) \leq x_1^b(t) \quad \text{for t } \in [s, T_1] \text{ such that } x_1^a(t) > z_1^m.$$

Thus trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ satisfy conditions of Theorem 2 of the appendix from which we get

$$\Delta_b^a E_x[\int_0^{t_1} c^m(X_1(t))dt] \leq 0$$

Since this is true for every renewal cycle when Trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ are different, we get,

$$\lim_{T\to\infty} \frac{1}{T} E_x^a[\int_0^T c^m(X_1(t))dt] \leq \lim_{T\to\infty} \frac{1}{T} E_x^b[\int_0^T c^m(X_1(t))dt]$$

Case 2: $p_1 \leq z_1^m$. There are two subcases corresponding to whether the switching curve intersects the line $x_1 = z_1^m$ or not.

Subcase 1: The switching curve does not intersect the line $x_1 = z_1^m$.

The policy a is now chosen such that the hedging point is $\boldsymbol{z}^a = (z_1^m, z_2)$, and the switching curve is defined by the straight line $x_1^a = z_1^m, x_2 < z_1^m$ (see Figure 4). Consider two trajectories $\boldsymbol{X}^a$, $\boldsymbol{X}^b$ which evolve respectively under policies a and b, and originate at their respective hedging point. From Corollary 2, $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ leave and reach these hedging points at the same instants. Note the switching curve for policy b does not intersect the line $x_1 = z_1^m$, and since $z_1 < z_1^m$ and $z_2 = z_2^m$, trajectories $\boldsymbol{X}^a$, $\boldsymbol{X}^b$ are such that Note that the Policy a is a static priority policy that always produces Part Type 1 first. Since initially $X_1^a(0) > X_1^b(0)$ and $\boldsymbol{X}_2^a(0) = \boldsymbol{X}_2^b(0)$, the sample path comparison yields,

$$X_2^a(t) \leq X_2^b(t) \tag{16}$$

$$X_1^b(t) \leq X_1^a(t) \tag{17}$$

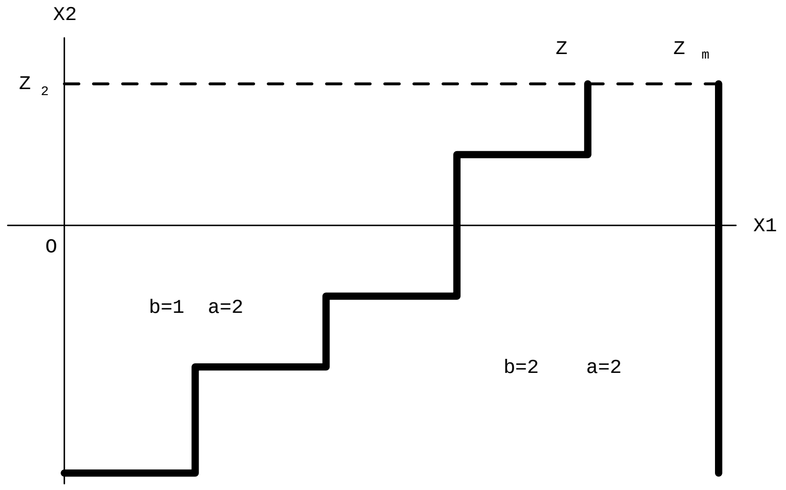

Figure 4: Case 2 $p_1 \leq z_1^m$

Using (17), and disecting the problem as did Srivatsan and Dallery (1998) in three cases (when $x_2^b(t)$ and $x_2^a(t)$ are both positive, when $x_2^b(t)$ is positive and $x_2^a(t)$ is negative, and when both are positive), a straightforward computation reveals that for all t,

$$\Delta_b^a c(X_1(t), X_2(t)) \leq \Delta_b^a [c_1(X_1(t)) - b_2 X_2(t)]$$

Furthermore,

$$\begin{aligned} \Delta_b^a [c_1(X_1(t)) - b_2 X_2(t)] &= \Delta_b^a [c_1(X_1(t)) + \frac{\mu_2}{\mu_1} b_2 X_1(t) - b_2 \mu_2 W(t)] && (18) \\ &= \Delta_b^a [c^m(X_1(t))] - b_2 \mu_2 \Delta_b^a [W(t)] && (19) \end{aligned}$$

We have $\Delta_b^a W(0) > 0$. So, from Lemma 4 and (18) it follows that

$$\Delta_b^a E[c(\boldsymbol{X}(t))] \leq \Delta_b^a E_x[c^m(X_1(t))]$$

Under policy a, part Type 1 behaves as if it were following the optimal hedging point policy for the single-part-system where the arrival rate, the production rate and the function of cost are respectively, λ_1, μ_1 and c^m (see the appendix).

So we get

$$\Delta_b^a E_x[\int_0^T c^m(X_1(t)) dt] \leq 0$$

Since this is true for every renewal cycle when Trajectories $\boldsymbol{X}^a$ and $\boldsymbol{X}^b$ leave their hedging point, we get,

$$\lim_{T\to\infty} \frac{1}{T} E_x^a[\int_0^T c^m(X_1(t))dt] \leq \lim_{T\to\infty} \frac{1}{T} E_x^b[\int_0^T c^m(X_1(t))dt].$$

Subcase 2: The switching curve does intersect the line $x_1 = z_1^m$ in the point $\hat{z} = (z_1^m, \hat{z}_2)$. Let Policy a be a hedging point policy with the same hedging point z as policy b and whose switching curve is the vertical line $x_1 = z_1^m$ for $x_2 < \hat{z}_2$ and the same as policy b elsewhere.

Consider two trajectories $\boldsymbol{X}^a$, $\boldsymbol{X}^b$ which evolve respectively under policy a and b, and that originate at the same point x at time $t = 0$. From Corollary 1, the trajectories pass through the point $\hat{\boldsymbol{z}}$ at the same time T_1. Using the same arguments as in the proof of subcase 1, it follow that for all t in $[0, T_1]$

$$\Delta_b^a E[c(\boldsymbol{X}(t))] \leq \Delta_b^a E[c_1(X_1(t)) - b_2 X_2(t)] \leq \Delta_b^a E[c^m(X_1(t))]]$$

Furthermore as in case 1, the trajectories of part type 1 under policies a and b satisfy the conditions of Theorem 2 of the appendix, and

$$\Delta_b^a E_x[\int_0^T c^m(X_1(t))dt] \leq 0$$

Hence the result. $\square$

Remark: In the case where $\mu_1 = \mu_2$, we actually have in case 1 $\Delta_b^a c(\boldsymbol{X}(t)) = \Delta_b^a c^m(X_1(t))$ for t in $[s, T_1]$ on the coupled path.

A.2 Results in the Single-Part-Type System

In this section, we derive results based on comparisons of the cost function of (1) for trajectories generated by two different policies over a given interval of time. The system considered here is a single-part-type system with exponential production and demand interarrival times, with rates μ and λ respectively. In this case, the controls associated with the policies are of the form:

$$C_a = \begin{cases} 0 & \text{the action is to idle} \\ 1 & \text{when the action is to produce} \end{cases}$$

For this system a hedging point policy is optimal. Following Buzacott and Shanthikumar (93), the optimal hedging point z is given by:

$$x_1 = z_1^m = \lfloor \frac{\ln\left(\frac{h}{b+h}\right)}{\ln \frac{\lambda}{\mu}} \rfloor \tag{20}$$

In most of the cases considered below, neither of the two policies is a hedging point policy. But the trajectories generated by these policies satisfy conditions that are relevant in the context of the two-part-type system. This policies are admissible scheduling policies in the sense of Veatch and Wein (1996): they are functions of the state X and the time instant t, and are nonanticipating with respect to X.

Let the expected cost of policy a in the interval $[s, f]$ be denoted as

$$J^a(s, f) = E\{\int_s^f c(x^a(t))dt | x^a(s) = x\}$$

(For notational simplicity, we do not explicitly express the dependence of this cost on the initial conditions $x^a(s)$.)
The difference in the expected costs of policies a and b in the interval $[s, f]$ is represented by

$$\Delta J_b^a(s, f) = J^a(s, f) - J^b(s, f)$$

Let J^b be the long term average cost of policy b . Let J^* be the optimal long average cost.

We first show that for some specified initial conditions and the same final conditions, a trajectory generated by a hedging point policy in an interval $[0, T]$ incurs a cost no greater than that generated by any other feasible policy. Note that our result pertains to total costs incurred over a finite interval.

Lemma 6 *Consider two trajectories X^a and X^b and an interval of time $[0, T]$ such that: X^a is generated by the optimal hedging point policy and X^b is generated by some other policy over $[0, T]$ and:*

1) $X^b(0) \leq X^a(0) \leq z$ or $X^b(0) \geq X^a(0) \geq z$;

2) $X^a(T) = X^b(T) = z$.

Then, $\Delta J^{ab}(0, T) \leq 0$.

Proof: Using the uniformization of Markovian processes, we can take λ and μ such that $\lambda + \mu = 1$ without loss of generality. Let G be the infinitesimal generator of the controlled Markov process with Policy π (with $\pi = b$ or $\pi = a$),

$$G(\pi)f(x) = \lim_{h \to 0} \frac{E[f(x(h)) - f(x(0))]}{h}$$

where f is a function of the states. Using Dynkin's formula (see Karlin and Taylor (81)) in the interval $[0, T]$, we have:

$$E[\int_0^T G(\pi)f(X)dt] = E[f(x(T))] - f(x(0)) = f(z) - f(x(0)) \tag{21}$$

There exists a relative value function V which verifies the following dynamic equation:

$$J^\pi - c(x) = G(\pi)V(x) \tag{22}$$

Since a is the optimal policy, It follows from (22) that,

$$\begin{aligned} G(a)V(x) &= J^* - c(x) \\ &= \min_\pi(J^\pi) - c(x) \\ &= \min_\pi(G(\pi)V(x)) \end{aligned}$$

Thus, we have :

$$G(b)V(x) \geq J^* - c(x) \tag{23}$$

Integrating and taking the expected value, we obtain from (23):

$$\begin{aligned} E(\int_0^T G(b)V(X(t))dt) &\geq J^*T - E[\int_0^T c(X(t))dt] \\ &\geq J^*T - J^b(0,T) \qquad (24) \\ E(\int_0^T G(a)V(X(t))dt) &= J^*T - J^a(0,T) \qquad (25) \end{aligned}$$

Combining (21) and (24) it follows,

$$J^b(0,T) \geq J^* + V(X^b(0)) - V(z).$$

From condition 3 of the lemma, we have $V(x^b(0)) \geq V(x^a(0))$. Thus, from (25)

$$J^b(0,T) \geq J^a + V(X^a(0)) - V(z) = J^a(0,T)$$

□

Next, we establish a similar result for the case where the initial conditions of the two trajectories satisfy a stricter condition than in Lemma 6 but the final conditions of the trajectories are not as restrictive as in Lemma 6.

Lemma 7 *Consider two trajectories X^a and X^b and an interval $[0,T]$ such X^a is generated by the optimal hedging point policy and X^b by some other policy over $[0,T]$ and:*

1) $X^b(0) \geq X^a(0) \geq z$;

2) $X^b(T) \geq X^a(T) = z$.

Then, $\Delta J^{ab}(0,T) \leq 0$.

Proof: There are two cases to be considered.

Case 1: Trajectory X^b does not go below z in $[0,T]$. Let t_m a time instant when Trajectory X^b takes its minimum value in $[0,T]$. Define a new trajectory X^c as follows:

$$X^c = \begin{cases} X^b(t) - (X^b(t_m) - z) & 0 \leq t \leq t_m \\ z & t_m < t \leq T \end{cases}$$

By definition, for all t in $[0,T]$

$$X^b(t) \geq X^c(t) \geq z$$

which gives, $J^b(0,T) \geq J^c(0,T)$. The Trajectories X^a and X^c satisfy the conditions of Lemma 6. Thus,

$$J^b(T) \geq J^c(T) \geq J^a(T)$$

Case 2: $X^b(t) < z$ for some t, $0 < t < T$. Let t_n be the last time instant such that X^b crosses z from below. From condition 2 of the lemma, t_n does exist. Define a new trajectory X^c as follows:

$$X^c = \begin{cases} X^b(t) - (X^b(t_m) - z) & 0 \leq t \leq t_n \\ z & t_n < t \leq T \end{cases}$$

Similarly as in the case 1, we obtain $J^b(t) \geq J^c(t) \geq J^a(t)$. □

We now combine the results of Lemmas 6 and 7 to show that if we choose trajectories X^a and X^b and an interval $[0,T]$ such that X^a, X^b satisfy specific conditions over an interval $[0, T]$, have the same initial and final conditions, and their final values at T are greater than the optimal hedging point value, then the cost incurred by X^a is no greater than that incurred by X^b.

Theorem 2 *Consider two trajectories X^a and X^b and an interval $[0, T]$ such that X^a and X^b satisfy Condition* **C1** *over the interval and:*

1) X^a is generated by the optimal hedging point policy for all $t \in [0, T]$ such that $X^a(t) < z$;

2) $X^a(t) \leq X^b(t)$ for all $t \in [0, T]$ such that $X^a(t) > z$.

3)$X^a(0) = X^b(0)$;

4)$X^b(T) = X^a(T) \geq z$. Then, $\Delta J^{ab}(0, T) \leq 0$

Proof: We have to consider two cases.
Case 1: $X^a(0) = X^b(0) \geq z$. Now, if $X^a(t) \geq z$ for all t in $[0, T]$, then from Condition 2, $hz \leq c(X^a(t)) \leq c(X^b(t))$ for all t in $[0, T]$.

The more interesting case is when Trajectory X^a goes below z at least once in $[0, T]$. Define t_i the time instant when Trajectory X^a crosses z from above, for the ith time. From Condition 4, there also exists at least one time instant such that Trajectory X^a crosses z from below. Define τ_i the time instant when Trajectory X^a crosses z from below, for the ith time. By definition we have:

$$\Delta_b^a J(0, T) = \Delta_b^a J(0, t_1) + \sum_{i=1}^{n-1} [\Delta_b^a J(t_i, \tau_i) + \Delta_b^a J(\tau_i, t_i)] + \Delta_b^a J(t_n, \tau_n) + \Delta_b^a J(\tau_n, T) \quad (26)$$

For all t in $[\tau_i, t_i]$ and in $[\tau_n, T]$, we obtain as above $hz \leq c(X^a(t)) \leq c(X^b(t))$.
For all t in $[t_i, \tau_i]$ and in $[0, t_1]$, Trajectories X^a and X^b satisfy conditions of Lemma 7. Thus, we have:

$$\Delta_b^a J(t_i, \tau_i) \leq 0 \text{ and } \Delta_b^a J(0, t_1) \leq 0.$$

Hence, from (26) we obtain $\Delta_b^a J(0, T) \leq 0$.
Case 2: $X^a(0) = X^b(0) < z$. From the conditions of the lemma, there exists a time instant such that X^a reaches z. Let τ be the first of these instants. In the interval $[0, \tau]$, the trajectories satisfy the conditions of the lemma 6. Therefore, $\Delta_b^a J[0, \tau] \leq 0$. In $[\tau, T]$, we can use the same arguments as in Case 1. Thus we have:

$$\Delta_b^a J[0, T] = \Delta_b^a J[0, \tau] + \Delta_b^a J[\tau, T] \leq 0$$

□