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

FRANCIS DE VERICOURT

Laboratoire d'Informatique de Paris 6 (LIP6-CNRS), Universite Pierre et Marie Curie, Paris, France, francis.vericourt@lip6.fr

FIKRI KARAESMEN and YVES DALLERY

Laboratoire Productique Logistique, Ecole Centrale Paris, Chatenay-malabry, France
fikri@pl.ecp.fr • dallery@pl.ecp.fr

(Received May 1998; revisions received November 1998, February 1999; accepted March 1999)

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 multiclass 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 that can be set 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.

The multiclass make-to-stock queueing problem was first considered by Zheng and Zipkin (1990), who showed 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 first-come first-served (FCFS) policy. These results were later generalized to multiple products by Zipkin (1995). Wein (1992) proposed a Brownian approximation for the multiclass 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 products 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, with 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 multiclass make-to-stock queueing problem (in further generality than proven) is a hedging point policy combined with monotone switching curves 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.

It is also interesting to note the similarities between make-to-stock queues and continuous flow models. For example, in the single product case, the solutions of the optimal control problems for the make-to-stock queue and continuous flow model of Bielecki and Kumar (1988) are closely related; in both cases the optimal policy is completely characterized by a single hedging point. In the multiproduct case, for the continuous flow two-part-type

Subject classifications: Production/scheduling, Stochastic: multi-item. Queues: make-to-stock system. Inventory/production: optimal policies.
Area of review: MANUFACTURING OPERATIONS.

0030-364X/00/4805-0811 $05.00
1526-5463 electronic ISSN

problem, Srivatsan and Dallery (1998) have recently provided a partial (but exact) characterization of the optimal hedging point policy. This exact characterization prompts the question of whether similar properties carry over to the conceptually related but considerably different case of the discrete part make-to-stock queue.

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 because of the passage from the continuous to the discrete case, but mainly because of 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 §2. In §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 §4. In §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 §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 or 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. Because 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. $\mathbf{X}(t) = (X_1(t), X_2(t))$ is then the state of the system. Let C_a be the control associated with a Markov policy a. We have

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

We consider a unit holding cost h_i and a unit backorder cost b_i per unit of time for part type i. In the state $\mathbf{X}$, the system incurs cost rate of $c(\mathbf{X}) = \Sigma_{i=1}^{2} c_i(X_i)$ where the individual part type costs c_i are

$$c_i(X_i(t)) = \begin{cases} h_i X_i(t) & X_i(t) \geqslant 0, \\ -b_i X_i(t) & X_i(t) \leqslant 0. \end{cases}$$

The objective is then to find the policy that minimizes the long-run average cost:

$$\min_a \limsup_{t\to\infty} \frac{1}{t} E_x^a \left[\int_0^t c(\mathbf{X}(t))\, dt \right], \tag{1}$$

where E_x^a denotes the conditional expectation given the control policy a and the initial condition $x = \mathbf{X}(0)$.

To solve the optimal control problem (1), a classical approach is to derive the dynamic programming optimality equations. Following Veatch and Wein (1996), with g^* the optimal average cost rate, $V(x)$ the relative value function, we have

$$\begin{aligned} V(x) + (g^*/\Lambda) &= 1/\Lambda[c(x) + \lambda_1 V(x_1 - 1, x_2) + \lambda_2 V(x_1, x_2 - 1) \\ &\quad + \mu V(x) + \min(0, \mu_1 \Delta_1 V(x), \mu_2 \Delta_2 V(x))] \end{aligned} \tag{2}$$

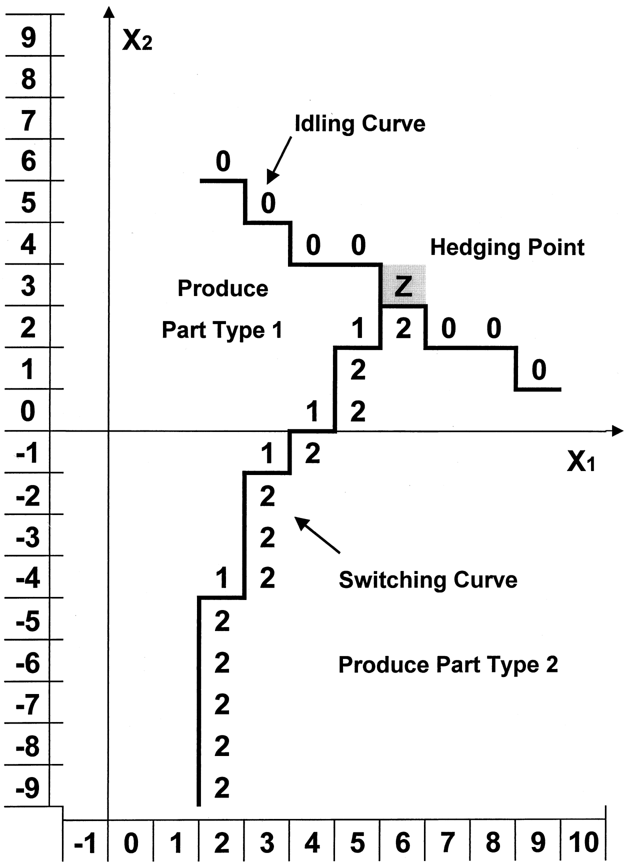
where

$$\Delta_1 V(x) = V(x_1 + 1, x_2) - V(x),$$
$$\Delta_2 V(x) = V(x_1, x_2 + 1) - V(x),$$
$$\mu = \max(\mu_1, \mu_2), \qquad \Lambda = \lambda_1 + \lambda_2 + \mu.$$

2.2. Optimal Control and Hedging Point Policies

The optimality equation (2) 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).

Figure 1. Hedging point policy.



Previous work on this problem supports this conjecture. First, 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. We thereby restrict our attention to this class of policies hereafter.

To give a more precise definition of *hedging point policies* for the manufacturing system in §2.1, it will be useful to define the class of monotone policies as in Veatch and Wein (1996). Let $\mathscr{I}$ be the set of states in which the control policy idles the machine. Furthermore, to separate resource sharing from idling, let $\bar{\mathscr{B}}_i$ denote the set of states in which the preferred part type for production is part i regardless of the idling decision (i.e., $\bar{\mathscr{B}}_i = \{x : \min_{j=1,2}\{\mu_j \Delta_j V(x)\} = \mu_i \Delta_i V(x)\}$). Then $\bar{\mathscr{B}}_i = \{x : x_1 < s_B(x_2)\}$ for an increasing curve $s_B(x_2)$ where $-\infty \leqslant s_B(x_2) \leqslant \infty$ and $\bar{\mathscr{I}} = \{x :\ x_1 < s_I(x_2)\}$ for some decreasing curve, $s_I(x_2)$, where $0 \leqslant s_I(x_2) \leqslant \infty$.

A hedging point policy is then a monotone policy whose stationary behavior is entirely characterized by the hedging point (z_1, z_2) (the intersection of the curves s_B and s_I) and the portion of the switching curves $s_B(x_2)$ with $x_2 \leqslant z_2$.

In the rest of the paper we concentrate on hedging point policies with starting points in the region $x_1 \leqslant z_1$ and $x_2 \leqslant z_2$ because other initial conditions do not affect the stationary behavior of the system.

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 1. 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 the control action at that time. Specifically, we have

$$C_a = \begin{cases} 0 & \text{Idle: The work is not used;} \\ 1 & \text{Produce type 1: The work is allocated to type 1;} \\ 2 & \text{Produce type 2: The work is allocated to type 2.} \end{cases}$$

3. When $C_a(t) \neq 0$, the allocation of the "work" to a given part type may result in the "instantaneous 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 $p_i = \mu_i/\mu$, 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 activities (times)" instead of "production activities (times)." Note that not all service activity completions correspond to actual production completion of a part.

An intuitive interpretation of this model is that the machine is not perfectly reliable: It can produce parts that 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.

Note that the EQ Model corresponds to a uniformization of the service processes in the original model. This simple transformation facilitates sample path comparisons because the service times after the transformation become independent of the policy. The proof of the above property is given here for clarity and to introduce some useful notation to be used in the sequel.

Proposition 1. *Under the same control policy, the EQ model follows the same probability law as the original model.*

Proof. Consider first the machine and the two 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: $P(\{U_i=1\})=p_i$ 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. Consider any state (x_1, x_2) and the corresponding control $C_a(x_1, x_2)$. Let $C_a(x_1, x_2)=1$ $(=2)$ without loss of generality. In the original model the transition rate to state (x_1+1, x_2) $((x_1, x_2+1))$ is μ_1 (μ_2). In the EQ model, the production transition rate for part type i under the identical control is given by $\mu p_i = \mu_i$. Therefore the production transition rates are equal for both models for the same state and control. Because the arrival rates in both models are also identical for all states, the original model and the EQ model have the same probability law. □

3. SOME PROPERTIES OF HEDGING POINT POLICIES

This section presents some properties of hedging point policies. They will constitute the basis of the proof for the characterization of optimal hedging point policies. These results, as others in this paper, are based 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.

More precisely, consider two trajectories $\mathbf{X}^a$ and $\mathbf{X}^b$ generated by two policies a and b. Consider the EQ model and T and U_i as defined 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 as follows: 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 and $U_2, u_1^1, u_1^2, \ldots, u_1^n, \ldots$, and $u_2^1, u_2^2, \ldots, u_2^n, \ldots$, where $u_i^n = 0$ or 1. These realizations are common for both policies. Let T^n be the time instant of the end of the nth service activity. $T^n = \Sigma_{k=1}^n t^k$, and T^n is the same for both policies. At instant T^n, the nature of the new event is then given by the realization of random variable U_i where $i=1$ or $i=2$ depending on the choice stated by the policy. Let $n_i^a(t) = \Sigma_{k:T^k<t} u_i^k$ be the number of times that policy a has chosen i before time t and similarly for policy b. If policy a chooses part type i, n^{a_i} is incremented but n^{a_j} is not.

For the development in the sequel, it will also be useful to couple the demand arrival processes as well as the service activities. In this case, we denote by $\tilde{T}^n$ the time instant of the nth event, which can either be a demand arrival or a service activity, and we denote by $\tilde{n}(t)$ the associated counting process. Since $\tilde{T}^n$ is constructed by superposing the arrival time sequences with the sequence T^n, it is independent of the control policy.

Following this approach, Lemma 1 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 instants of the batches must be identical. Thus, under certain conditions the lemma provides us the positions of coupled trajectories at the same instant 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 $\mathbf{X}^a$, $\mathbf{X}^b$ 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 $i=2$, $X_i^a(t) \leqslant X_i^a(T_a)$ (respectively $X_i^b(t) \leqslant X_i^b(T_b)$) and T_a (resp. T_b) is the first time that $\mathbf{X}^a$ (resp. $\mathbf{X}^b$) reaches $\mathbf{X}^a(T_a)$ (resp $\mathbf{X}^b(T_b)$);*
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. We denote by $n(t)$ the total number of service activities up to (and including) time instant t. From the definition of n_i^a, and because 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 u_i^k 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 Lemma 1, it follows that for $i=1$ or $i=2$,

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

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

$$X_1^a(T_a) - X_1^a(0) > X_1^b(T_a) - X_1^b(0) \quad \text{and}$$

$$X_2^a(T_a) - X_2^a(0) \geqslant X_2^b(T_a) - X_2^b(0). \tag{7}$$

By combining (7) with (4) we obtain that $n_1^a(T_a) > n_1^b(T_a)$ and $n_2^a(T_a) \geqslant n_2^b(T_a)$, which from (3) is impossible. Using similar arguments when $T_a < T_b$, we have $T_a = T_b$. □

Lemma 1 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 hedging point policies in the following corollaries.

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

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

2. $\mathbf{X}^a$ *and* $\mathbf{X}^b$ *start at the same initial point, such that* $X_1(0) \leqslant \hat{z}_1$ *and* $X_2(0) \leqslant \hat{z}_2$, *then the trajectories reach* $\hat{z}$ *at the same time instant.*

Corollary 2. *A hedging point policy a is stable if* $(\lambda_1/\mu_1 + \lambda_2/\mu_2) < 1$.

The proofs of these corollaries can be found in de Véricourt et al. (1998).

4. PARTIAL CHARACTERIZATION OF THE OPTIMAL HEDGING POINT POLICY

Consider the two-part-type system introduced in §2.1. Without loss of generality, let the two-part-types be numbered such that $b_1\mu_1 \geqslant b_2\mu_2$. We derive in this section a structural result for the optimal hedging point policy of the two-part-type system introduced in §2.1. The main idea is to relate the optimal control problem of the two-part-type system to a single-part-type problem. Informally, we will exploit the fact that the instantaneous cost function can be expressed as $c(x_1, x_2) = c^m(x_1) - f(x_1, x_2)$ for $x_2 \leqslant 0$, 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

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

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

Definition 2. Let the aggregate workload be defined as $W(t) = X_1(t)/\mu_1 + X_2(t)/\mu_2$.

We denote by $W^a(t)$, the aggregate workload under policy a. We also use the following notations: $W(\mathbf{X}) = W(\mathbf{X}(t)) = W(t)$. 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 expressed 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 2. *Consider 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* $\Delta_b^a E[W(t)] = E[W^a(t)] - E[W^b(t)] = W^a(0) - W^b(0)$.

Proof. Consider the EQ model. Let $\tilde{T}^n$ be the time of the nth event and $\tilde{n}(t)$ the associated counting process, as defined in §3. $\tilde{T}^n$ can be a demand arrival or a service activity (recall that in the EQ model, a service activity does not necessarily correspond to a production completion). Consider a time instant t, we have $W(t) = W(\tilde{T}^n)$ with $n = \tilde{n}(t)$.

Suppose then that the instant $\tilde{T}^n$ corresponds to a service activity. It follows for policy a when $C_a(\tilde{T}^n) = i$ that

$$E[W^a(\tilde{T}^n)] = E\left|W^a(\tilde{T}^{n-1}) + \frac{U_i}{\mu_i}\right|, \tag{8}$$

where U_i is the discrete random variable of the EQ model. From the definition of U_i we obtain that $E[U_i/\mu_i] = E[U_i]/\mu_i = 1/\mu$. Note that this value does not depend on the part type. Thus, from (8) we have

$$E[W^a(\tilde{T}^n)] = E[W^a(\tilde{T}^{n-1})] + \frac{1}{\mu}. \tag{9}$$

If $\tilde{T}^n$ corresponds to an arrival of type i, then

$$E[W^a(\tilde{T}^n)] = E[W^a(\tilde{T}^{n-1})] - \frac{1}{\mu_i}. \tag{10}$$

Similarly, for policy b,

$$E[W^b(\tilde{T}^n)] = E[W^b(\tilde{T}^{n-1})] - \frac{1}{\mu} \quad \text{or}$$

$$E[W^b(\tilde{T}^n)] = E[W^b(\tilde{T}^{n-1})] - \frac{1}{\mu_i}. \tag{11}$$

As presented in §3, because of the coupling, arrival and service activity instants do not depend on the control policy. Hence, $\tilde{T}^i$'s are the same under policies a and b. From Equations (9), (10), and (11), it follows then that $\Delta_b^a E[W(\tilde{T}^n)] = \Delta_b^a E[W(\tilde{T}^{n-1})]$ giving us the desired result. □

Based on Lemma 1, 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* $b_1\mu_1 \geqslant b_2\mu_2$. *When* $x_2 < 0$, *the switching curve of the optimal hedging point policy for this system is the straight-line defined by*

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

Table 1. 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	1	10	0.25	200	5	5
4	1.2	0.1	2	1	1	1	50	20	3

and the optimal hedging point policy a is of the form

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

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:

Corollary 3. $z_1^m = 0 \iff h_1\mu_1 - h_1\lambda_1 > b_1\lambda_1 - b_2\mu_2$.

5. NUMERICAL RESULTS AND HEURISTIC POLICIES

5.1. Numerical Results

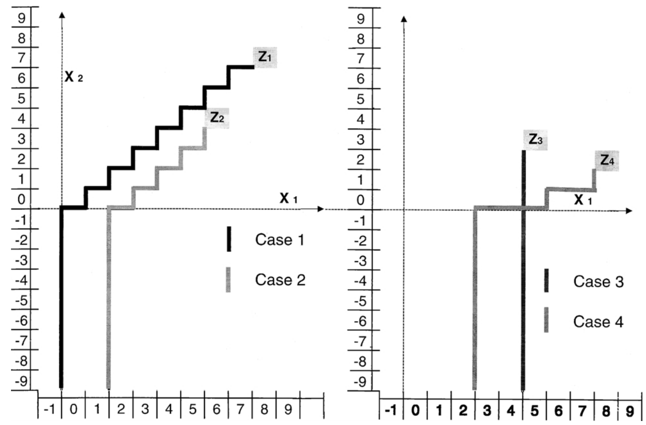
To numerically compute optimal policies, we have used the value iteration algorithm to solve the optimality equation (2) in a truncated state space. All the numerical examples we have tested confirm the optimality of a hedging point policy (when $\mu_1 \neq \mu_2$) as well as the results in Theorem 1.

We present a small subset of the computational experiments 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 1.

Figure 2 displays the optimal switching curves for the four cases of Table 1. It can be seen that the position of the switching curve (when $x_2 \leqslant 0$) as computed by the value iteration algorithm is equal to the theoretical value given by Theorem 1, which is reported as Z^m in Table 1.

Theorem 1 enables a quick interpretation of the optimal policies displayed in Figure 2. Cases 1 and 2 differ only in their backlog costs. The respective position of the vertical line reflects the differences in ratios b_1/h_1, b_2/h_1 for these two cases. Case 3 is an example where part 1 is much more important with respect to part 2 in holding and backlog costs and demand rates; the optimal policy is in this case is a static priority policy. Finally, in case 4, the two parts are asymmetric with respect to demand and service rates as well as backlog costs again leading to a vertical line away from zero.

Figure 2. Numerical results.



As seen in the above examples, the position of the optimal switching curve when $x_2 < 0$ partially reflects the asymmetry of costs. Another determining factor for the position of the line that emerges from Theorem 1 is ρ_1, the traffic load in isolation of part 1. Z^m is increasing in ρ_1 and can take large values when part 1 dominates part 2 in terms of its traffic load (even when costs may be almost symmetric).

5.2. Myopic Allocation

In §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") that 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, myopic allocation is based on intutive but informal arguments. After its short presentation, we give a partial justification of myopic allocation using Theorem 1.

The main idea of myopic allocation is to look ahead a service time of part-type i, say S_i. The policy then allocates the production capacity to the part type that increases the expected instantaneous cost (due to part-type i) 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, the expected instantaneous cost (of part i) after the completion of the service time is given by $g(x_i) = E[c(x_i + 1 - D(S_i))]$. Thus, $\mu_i \Delta g(x_i) = g(x_i + 1) - g(x_i)$ is the rate at which producing part-type i increases the instantaneous expected cost due to part-type i. The myopic heuristic then selects the part-type i with the minimum $\mu_i \Delta g(x_i)$. As an improvement of this starting idea Peña-Perez and Zipkin (1997) then suggest using the cost function $g(x_i) = E[c(x_i + 1 - D(T_i))]$ where T_i is exponentially distributed with rate $(1 - \rho_i)\mu_i$. The rationale behind the choice of T_i is that to increase x_i by 1, the machine will not only produce one unit of i but will also respond to the new demands until the inventory level reaches $x_i + 1$. This "replenishment time" has a mean proportional to $(1 - \rho_i)^{-1}$. The improvement brought by the incorporation of the replenishment time as the look-ahead period was verified by their numerical 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. The following lemma provides a formal justification of the excellent performance of these myopic approaches.

Lemma 3. *The myopic allocation policy applied with the sojourn time T_i is an optimal policy when $x_2 < 0$.*

Proof. A direct calculation of the allocation index with T_i yields that the myopic allocation policy states to produce part-type-1 when $x_2 < 0$ if and only if

$$x_1 < \min\left\{x_1 : \rho_1^{x_1+1} > \frac{\mu_1 h_1 + \mu_2 b_2}{\mu_1 h_1 + \mu_1 b_1}\right\}$$

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

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

This lemma explains the good performance of the myopic policy computed with the sojourn time. Another direct calculation shows that, for the myopic policy with the service time, the switching curve is also a straight line when $x_2 < 0$ with

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

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 conditions 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 comparison 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 (nonnegative) 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 product and can, in certain cases, be significantly large, depending on the cost and traffic parameters.

Similar results have been shown for an analogous continuous flow model. On 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, the fact that the optimal policy has the same structure for two models that represent randomness in very different ways indicates 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 multiproduct system. These results could also be useful for the multiproduct multistage 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.

APPENDIX

RESULTS IN THE SINGLE-PART-TYPE SYSTEM

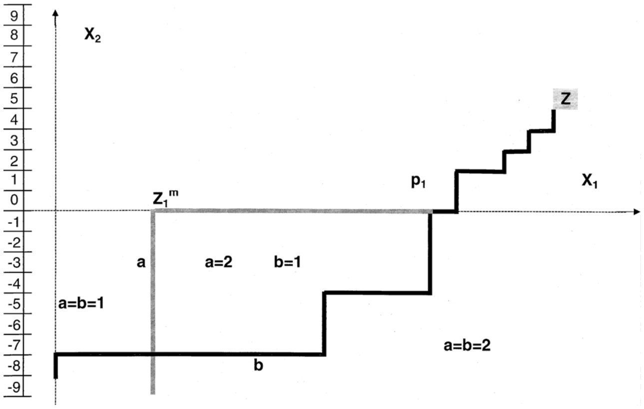
We now derive a result based on comparisons of the cost function of (1) for trajectories generated by two different policies over a given interval of time. The system considered here is a single-part-type system with exponential production and demand interarrival times, with rates μ and λ, respectively. 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. As noted by Buzacott and Shanthikumar (1993) and Veatch and Wein (1996), the optimal hedging point z is given by $x_1 = z_1^m = \lfloor \ln(h/(b+h))/\ln(\lambda/\mu) \rfloor$.

Theorem 2 shows that if we choose trajectories X^a and X^b such that they satisfy specific conditions on 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.

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



THEOREM 2. *Consider two trajectories X^a and X^b and an interval $[0,T]$ such that X^a and X^b satisfy the following conditions*:

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

(2) *For all $t \in [0,T]$ such that $X^a(t) \geqslant z$, $X^a(t) \leqslant X^b(t)$*;

(3) *$X^a(0) = X^b(0)$ and $X^b(T) = X^a(T) \geqslant z$.*

Then, $E[\int_0^T c(x^a(T))\,dt \mid x^a(0)=x] \leqslant E[\int_0^T c(x^b(t))\,dt \mid x^b(0)=x]$.

PROOF. The proof can be found in de Véricourt et al. (1998) and is an adaptation of a corresponding result by Srivatsan and Dallery (1998).

Note that the trajectories considered are not generated by hedging point policies. But they satisfy conditions that are relevant in the context of the two-part-type system as presented in the following proof.

Proof of Theorem 1

We prove the theorem by contradiction. Let Policy b be an optimal hedging point policy with a nonnegative hedging point, $\mathbf{z} = (z_1, z_2)$, and a switching curve, which differs from the straight 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. For each case, we construct a policy a which is better than b. Here, we give the proof of the first case where we suppose that $p_1 > z_1^m$. The reader is referred to de Véricourt et al. (1998) for a proof of the second case.

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 $x_2 = 0$, $z_1^m \leqslant x_1 \leqslant p_1$; the same as the switching curve of Policy b elsewhere. Policies a and b are illustrated in Figure 3. Consider two trajectories, $\mathbf{X}^a$ and $\mathbf{X}^b$, that start at the hedging point and evolve under Policies a and b, respectively. These trajectories can separate only in the region where the controls of policies a and b are different. Note that this region is included in the region $x_2 < 0$ and $x_1 < p_1$. Let s^- denote the 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.

When trajectories $\mathbf{X}^a$ and $\mathbf{X}^b$ separate again, the above scenario restarts. Thus, it is sufficient to prove that Policy a is better than Policy b for every renewal cycle. This is done in what follows.

Consider the costs of Policies a and b over the interval $[s, T_1]$. From §2.1, the instantaneous cost function for $x_2 < 0$, can be expressed as: $c(x_1, x_2) = c^m(x_1) - b_2\mu_2 W(x_1, x_2)$ for $x_2 \leqslant 0$, where W is the aggregate workload, and c^m is the instantaneous cost function given by

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

Because hedging point policies are stable, it follows that

$$\Delta_b^a E_x \left[\int_s^{T_1} c(\mathbf{X}(t))\,dt \right]$$
$$= \int_s^{T_1} \Delta_b^a E_x[c^m(X_1(t))]\,dt - b_2\mu_2 \int_s^{T_1} \Delta_b^a E_x[W(t)]\,dt.$$

At time s we have $W^a(s) = W^b(s)$, and for all t in $[s, T_1]$, the machine works at full capacity. Thus for t in $[s, T_1]$, $\mathbf{X}^a$ and $\mathbf{X}^b$ verify the conditions of Lemma 2, and we obtain $\Delta_b^a E_x[W(t)] = 0$. So the difference in expectations of cost between the two trajectories in $[s, T_1]$ is the same as that for system where the instantaneous cost function in the nonpositive 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 $X_2^a(t) < 0$ is the same as the one that is 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. Also, by construction, when $X_2^a(t) = 0$ we have $X_1^a(t) \leqslant X_1^b(t)$ for $t \in [s, T_1]$, such that $X_1^a(t) > z_1^m$ and the trajectories $\mathbf{X}^a$ and $\mathbf{X}^b$ satisfy conditions of Theorem 2, from which we have $\Delta_b^a E_x[\int_s^{T_1} c^m(X_1(t))\,dt] \leqslant 0$, giving us the result. □

ACKNOWLEDGMENTS

The authors are grateful to the area editor, Lawrence Wein, and the two referees, whose comments have led to a much improved version of the paper. They also thank one of the referees for pointing out several errors in a previous version.

REFERENCES

Bielecki, T., P. R. Kumar. 1988. Optimality of zero inventory policies for unreliable manufacturing systems. *Oper. Res.* **36** 532–546.

Buzacott, J. A., J. G. Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Ha, A. 1997. Optimal dynamic scheduling policy for a make-to-stock production system. *Oper. Res.* **45** 42–53.

Peña-Perez, A., P. Zipkin. 1997. Dynamic scheduling rules for a multiproduct make-to-stock queue. *Oper. Res.* **45** 919–930.

Srivatsan, N., Y. Dallery. 1998. Partial characterizations of optimal hedging point policies in unreliable two-part-type systems. *Oper. Res.* **46** 36–45.

Veatch, M., L. M. Wein. 1996. Scheduling a make-to-stock queue: Index policies and hedging points. *Oper. Res.* **44** 634–647.

de Vericourt, F., F. Karaesmen, Y. Dallery. 1998. Dynamic scheduling in a make-to-stock system: A partial characterisation of optimal policies. Technical Report 1998-22, Laboratoire d'Informatique de Paris 6.

Wein, L. M. 1992. Dynamic scheduling of a multiclass make-to-stock queue. *Oper. Res.* **40** 724–735.

Zheng, Y., P. Zipkin. 1990. A queueing model to analyze the value of centralized inventory information. *Oper. Res.* **38** 296–307.

Zipkin, P. 1995. Performance analysis of a multi-item production-inventory system under alternative policies. *Management Sci.* **41** 690–703.