OPTIMAL CONTROL OF A TWO-STATION TANDEM PRODUCTION/INVENTORY SYSTEM

Michael H. Veatch
Lawrence M. Wein

OR 264-92 February 1992

OPTIMAL CONTROL OF A TWO-STATION TANDEM PRODUCTION/INVENTORY SYSTEM

Michael H. Veatch
Operations Research Center, M.I.T.

Lawrence M. Wein
Sloan School of Management, M.I.T.

Abstract

A manufacturing facility consisting of two stations in tandem operates in a make-to-stock mode: after production, items are placed in a finished goods inventory that services an exogenous demand. Demand that cannot be met from inventory is back-ordered. Each station is modelled as a queue with controllable production rate, and the problem is to control these rates to minimize inventory holding and backordering costs. Optimal controls are computed using dynamic programming and compared with kanban and buffer control mechanisms, popular in manufacturing, and with the base stock mechanism popular in inventory/distribution systems. Conditions are found under which certain simple controls are optimal using stochastic coupling arguments. Insights are gained into when to hold work-in-process and finished goods inventory, comparable to previous studies of production lines in make-to-order and unlimited demand ("push") environments.

February 1992

dynamic programming models: control of tandem make-to-stock queues
multi-stage inventory/production: optimal control

Considerable attention has been given in recent years to viewing manufacturing facilities as production/inventory systems. This framework recognizes the importance not only of inventory control but also of queueing due to capacity constraints and uncertainty. In this paper, we consider a production/inventory system that consists of several stages in series that produce a single product. Each stage of the facility contains a single workstation that is modelled as a queue with controllable service rate. Once completed, items are counted as finished goods until they are consumed by an exogenous demand. As in inventory/distribution systems, production is driven by demand, but unlike standard inventory models there are capacity constraints and queueing in the production process. Demand that cannot be met from inventory is backordered and met by the next available finished item. Holding costs are incurred at each stage, as well as finished goods holding and backordering costs.

We consider the problem of finding optimal controls of the production rates for a long-run average or discounted cost criterion. In contrast, most studies of production line control problems assume that a relatively simple mechanism, such as buffers or kanbans, is used to control the system. Its performance is evaluated or the best policy using that mechanism is found. Restricting attention to a certain mechanism may be practical, given that optimizing a cost function is difficult and "optimal" policies may be impractical to implement; however, it is also desirable to know how the mechanisms compare to each other and to the optimal policy.

The demand environment of our production/inventory system is make-to-stock with complete backordering; other environments have been more widely studied. A make-to-order environment, where production of an item cannot begin until a demand is received, may be dictated by customization requirements on orders or chosen for economic reasons when customers will tolerate the waiting time. This environment corresponds to a tandem queue; its optimal control has been studied by Rosberg, Varaiya and Walrand (1982) and Weber and Stidham (1987), among others. A buffer control mechanism results in a tandem queue with blocking; lost sales can be incorporated in this model by including a finite first buffer. Approximate performance evaluation and optimal buffer placement have been studied; see Perros (1984), Smith and Daskalaki (1988), Hillier, Boling and So (1986), and the references therein. All of the environments described thus far are exogenous demand, or "pull" systems. Unlimited demand "push" systems can be modelled as closed queueing networks; buffer placement for these systems has been studied by Conway et al. (1988). Kanban policies, pioneered by Toyota (Sugimori et al. 1977), are studied by Mitra and Mitrani (1990) and Muckstadt and Tayur (1991) in an unlimited demand setting.

Several previous studies of multi-stage, single-product production/inventory systems have obtained approximate results for evaluating a specific control mechanism. Mitra and Mitrani (1991) and Cheng and Yao (1991) both study kanban policies; they also establish sample path and stochastic dominance of kanban mechanisms over traditional buffer mechanisms. The dominance is essentially due to the moveable buffers within a kanban cell, in contrast with the traditional fixed buffers. Base stock policies, motivated by distribution/inventory systems, are evaluated by Lee and Zipkin (1990) and Buzacott, Price and Shanthikumar (1991) using stage decomposition approximations. Base stock is not really a new concept in manufacturing since, as the second paper points out, MRP systems essentially use a base stock mechanism with a demand forecast included in the target stock levels. Constant work-in-process (CONWIP) can be viewed as a special case of this policy.

Our approach is to find optimal policies, using analytical and numerical methods, and compare them with some of the simpler control mechanisms being used in manufacturing. To accomplish this program, a simple two-station problem is considered. It is assumed that demand is Poisson, service times are exponential, and there are no set-up costs. It is hoped that the insights gained from this idealized system, with careful attention to its limitations, will be applicable to more realistic systems. It is encouraging to note that van Ryzin, Lou and Gershwin (1991) and Lou and van Ryzin (1989) obtained similar numerical results for a similar system in which the only source of uncertainty is unreliable machines, indicating that there is at least some robustness to our findings.

Two types of results are obtained in this paper. First, very simple policies are shown to be optimal under certain extreme conditions on the problem parameters. Veatch and Wein (1991) established that the optimal policies generally consist of a switching curve for each station, dividing the state space into an idle and busy region. We use stochastic coupling arguments to show that, for certain parameter values, these switching curves become essentially static priority rules. Conditions are found under which no inventory is held, as well as conditions under which all inventory is converted to finished goods (FG), i.e., the downstream station never idles unless it is starved. These results are comparable to those of Bielecki and Kumar (1988) for a single-stage production/inventory system. Interestingly, the popular base stock policy is shown to never be *exactly* optimal.

Second, numerical results are obtained using dynamic programming. These results further illustrate the tradeoffs of whether or not to hold inventory and whether to hold work-in-process (WIP) or FG. Holding WIP may seem to fly in the face of the just-in-time goal of eliminating WIP; in fact, our model provides a cost basis for deciding whether or not to hold WIP and FG. In a production/inventory system, WIP can perform two functions: not only does it serve as a buffer between asynchronous stations to increase throughput capacity (as in make-to-order systems), it can also supplement finished goods (FG) inventory to reduce backorders. The decision of where to place inventory depends on the relative holding costs and the rate at which WIP can be converted into FG. Less WIP is held when its holding cost is high, the utilization of the upstream station is low, or the discount rate is high.

Optimal policies are also compared to the best base stock, kanban, and fixed buffer policies. It is found that base stock policies are nearly optimal when the upstream station is heavily utilized and the discount rate is small or zero. Kanban policies outperform base stock policies when the downstream station is a bottleneck or discounting is significant. Fixed buffer policies are consistently the worst, though the degradation is not always significant. It is encouraging that base stock and kanban policies are within a few percent of optimal for most test cases, since a swithcing curve policy would be more difficult to implement. Every type of policy is sensitive to the stock levels or buffer sizes, so that obtaining accurate demand and production rate data and setting these levels correctly remains a very important issue.

The patterns that appear in the numerical study can only be extrapolated to more complex systems tentatively and qualitatively. It is reasonable to expect that the desirabilty of holding WIP would be similar for systems with more stations, probably with most WIP being held downstream. However, the amount of WIP held is always modulated by the service time variability, which is sometimes less in real systems than our exponential assumption. It also should be noted that we apply a cost to the average WIP; most studies of buffer allocation apply a cost or constraint to the maximum WIP, e.g., total buffer capacity (see McClain and Moodie 1991). As is well known, WIP can have other adverse effects than just the inventory holding cost (see, for example, Schonberger 1982).

Another result that may be of use in future research is a transformation of make-to-stock systems into make-to-order systems. This equivalence allows some of the methods developed for traditional tandem queues to be applied to make-to-stock systems. For example, approximate evaluation of stationary distributions for tandem queues with blocking can be used to quickly identify suboptimal policies for our system.

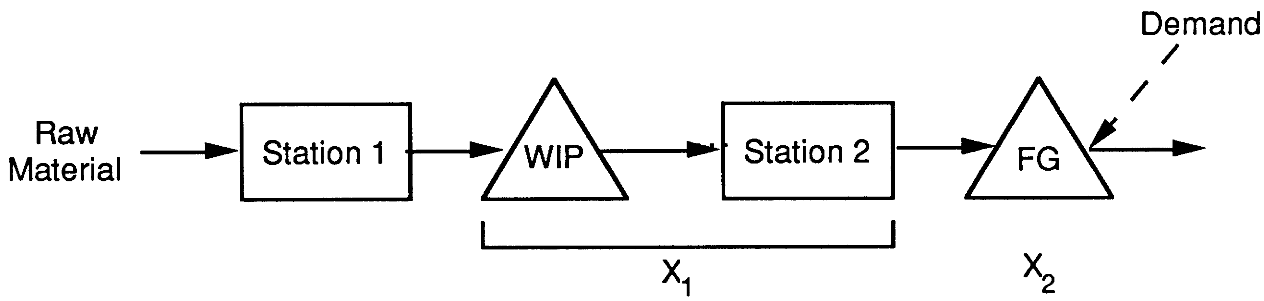

Figure 1: A Two-Stage Production/Inventory System

The remainder of the paper is organized as follows. The problem is formulated mathematically in Section 1 and several control mechanisms are defined in Section 2. The optimality of simple controls under special parameter values is proven in Section 3, with some of the proofs deferred to the Appendix. A connection with traditional make-to-order queues is made in Section 4 and dynamic programming numerical results are presented in Section 5.

1 Problem Desciption

Consider the two-stage tandem production system of Fig. 1. Jobs are released into the system, processed at stage 1, then held in a work-in-process (WIP) buffer. When released into stage 2, they are processed there and then placed in a finished goods (FG) inventory that services an exogenous demand. Demand that cannot be met from inventory is backordered and recorded as a negative inventory. Denote the system state at time t by $X(t) = (X_1(t), X_2(t))$, where X_1 is the number of jobs available for stage 2 processing (including any item being processed at stage 2) and X_2 is the FG inventory. Because the supply of raw material is unlimited, there is no queueing and no state variable at stage 1.

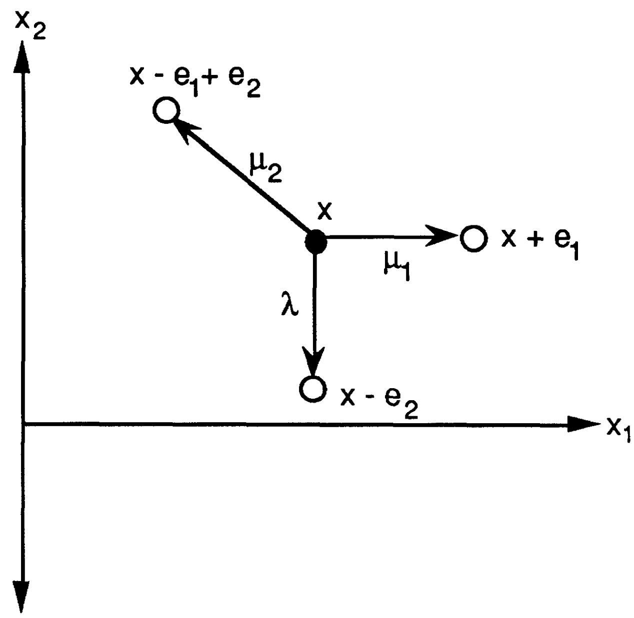

Figure 2: State Space Diagram

Stage i consists of a single machine that operates as a $\cdot/M/1$ queue with production rate μ_i controlled between 0 and $\overline{\mu}_i$. Associated with these controls are the transitions $x \rightarrow x + d_i$, where $d_1 = e_1$ for μ_1 and $d_2 = e_2 - e_1$ for μ_2, as shown in Fig. 2. Here e_i is the unit vector along the ith axis. Demands occur according to a Poisson process with rate λ and cause the transition $d_0 = -e_2$. Stability of the system requires that $\lambda < \overline{\mu}_i$ for $i = 1, 2$. An *admissible* control policy π is a function $\mu(X, t)$ that is nonanticipating, i.e., depends only on $\{X(s); s \leq t\}$, and obeys the control limits $0 \leq \mu_i \leq \overline{\mu}_i$ and $\mu_i(X, t) = 0$ if $X(t^-) + d_i \notin \mathbf{X} = \{x \in Z^2 : x_1 \geq 0\}$. Let Π denote the class of admissible policies. Because the system is memoryless, a Markov policy depending only on the current state x will be optimal; we denote this policy $\mu(x) = (\mu_1(x), \mu_2(x))$ for $x \in \mathbf{X}$.

The objective is to minimize WIP holding cost (incurred at a rate of one per job per unit time), FG holding cost $h > 1$, and FG backorder cost b, all discounted at a rate $\alpha > 0$ over an infinite time horizon. For policy π, the expected cost is

$$V^{\pi}(x) = E_x \int_0^{\infty} e^{-\alpha t} c(X(t))dt, \tag{1}$$

where $c(x) = x_1 + hx_2^+ + bx_2^-$. Here E_x denotes expectation given the initial state $X(0) = x$ and policy π. The optimal policy $\mu(x)$ achieves the minimum

$$V(x) = \min_{\pi \in \Pi} V^{\pi}(x) \tag{2}$$

simultaneously for all x. We will uniformize the process as in Lippman (1975) by defining the potential event rate $\Lambda = \overline{\mu}_1 + \overline{\mu}_2 + \lambda$. The n-stage cost function satisfies the dynamic programming equations

$$V_{n+1}(x) = \mathbf{T}V_n(x) \tag{3}$$

$$\begin{aligned} \mathbf{T}V(x) = \frac{1}{\Lambda + \alpha}[c(x) + \lambda V(x - e_2) + \overline{\mu}_1 \min\{V(x), V(x + e_1)\} \\ + \overline{\mu}_2 \min\{V(x), V(x - e_1 + e_2)\}], \end{aligned} \tag{4}$$

where we define $V_0(x) = 0$ and $V_n(x) = \infty$, $x \notin \mathbf{X}$. The infinite-horizon cost function satisfies

$$V(x) = \mathbf{T}V(x). \tag{5}$$

The form in which we have written (4) emphasizes that the optimal policy is *bang-bang*, i.e., $\mu_i(x) = 0$ or $\overline{\mu}_i$. Such a policy is specified by its idle and busy sets $\mathcal{I}_i = \{x \in \mathbf{X} : \mu_i(x) = 0\}$ and $\mathcal{B}_i = \mathbf{X} \setminus \mathcal{I}_i$. The existence of a Markov policy that achieves the minimum in (2) and the convergence of the n-stage policy and cost function to the infinite-horizon optimal policy and cost follow from the fact that only finitely many controls are considered at each state; see Bertsekas (1976).

An undiscounted, long-run average cost criterion will also be considered. In this

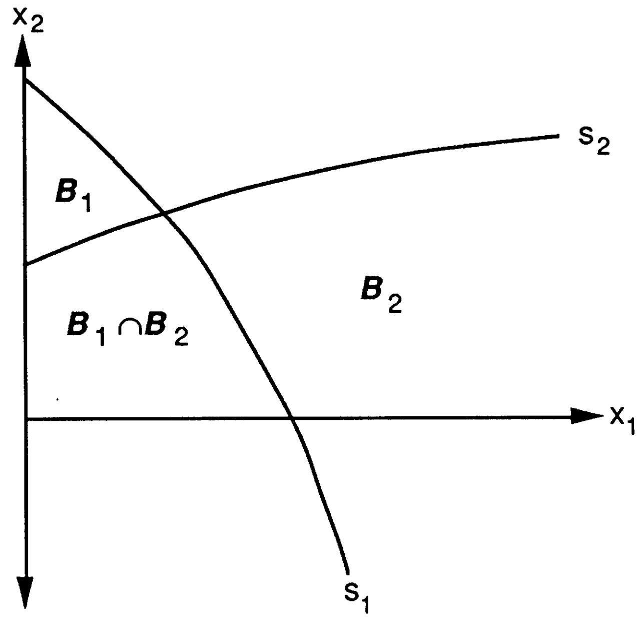

Figure 3: Properties of Optimal Policies

case the average cost per stage, g, and the relative cost of starting in state x, $V(x)$, satisfy

$$V(x) + g = \mathbf{T}V(x), \tag{6}$$

where we arbitrarily set $V(0,0) = 0$. Existence and convergence results can be obtained for (6) by letting $\alpha \to 0$ in (4) and exploiting the fact that there are only a finite number of "good" states; see Weber and Stidham (1987).

It is shown in Veatch and Wein (1991) that optimal policies have the following monotonicity property: there exist *switching functions* $s_i(x_1)$ such that $\mu_i(x) = 0$ if and only if $x_2 > s_i(x_1)$. Furthermore, these functions have derivatives (or more precisely, differences, since they are defined on Z^+) $s_1'(x_1) \leq -1$ and $s_2'(x_1) \geq 0$, as illustrated in Fig. 3.

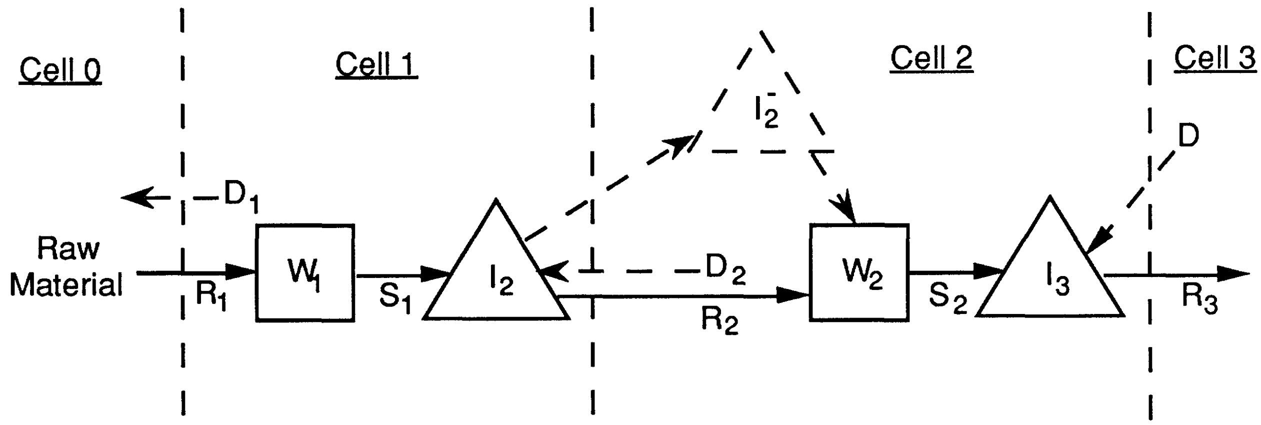

Figure 4: Material and Information Flow

2 Control Mechanisms

This section describes several control policies that have been studied and used in production lines. In order to describe the mechanism by which they are usually implemented, we begin with a more detailed model of material and information flow through the system. Associated with each station is a physical or organizational *cell.* As shown in Fig. 4, orders (hereafter called demands) are placed on cell 1 according to a policy that depends only on the state of cell 2. Either a release occurs immediately or the demand is backordered until there is inventory at the cell 1 output buffer. Define

$D_i(t)$ = demand placed by cell i on cell $i-1$ in $(0,t]$

$D(t)$ = exogenous demand in $(0,t]$

$R_i(t)$ = work released into cell i in $(0,t]$

$S_i(t)$ = service completions at station i in $(0,t]$

$W_i(t)$ = work at station i at time t

$I_i(t)$ = inventory position, cell i output buffer at time t.

The controls can be desribed in terms of demands as follows. Station i is idle when $W_i = 0$ and busy when $W_i > 0$; this defines $S_i(t)$. The dynamic equations are

$$\begin{aligned} R_1(t) &= D_1(t) && (7) \\ R_i(t) &= \min\{D_i(t), I_i(0) + S_{i-1}(t)\},\ i = 2,3 && (8) \\ I_i(t) &= I_i(0) + S_{i-1}(t) - D_i(t) && (9) \\ W_i(t) &= R_i(t) - S_i(t),\ i = 1,2. && (10) \end{aligned}$$

In the notation of section 1, WIP is $x_1 = I_2^+ + W_2$, FG is $x_2 = I_3$, and the dynamic equations are

$$\begin{aligned} x_1(t) &= x_1(0) + S_1(t) - S_2(t) && (11) \\ x_2(t) &= x_2(0) + S_2(t) - D(t). && (12) \end{aligned}$$

Base Stock

Under a base stock, one-for-one ordering policy each cell places an order upstream as soon as it receives an order. Hence, demands propagate through the system immediately and $D_i(t) = D(t)$. The application of this policy to production/inventory systems is discussed in Buzacott, Price and Shanthikumar (1991). Let c_1 and c_2 be the base stock levels for WIP and FG, respectively. The policy is characterized by the busy sets $\mathcal{B}_1 = \{x : x_1 + x_2 < c_1 + c_2\}$ and $\mathcal{B}_2 = \{x : x_1 > 0,\ x_2 < c_2\}$.

Kanban

A kanban policy has been applied to the make-to-stock environment by Mitra and Mitrani (1991). In terms of our model, the number of kanbans or cards in cell i is $c_i = I_i^- + W_i + I_{i+1}^+$; I_i^- represents the bulletin board and I_{i+1}^+ the output hopper in cell i. Demands occur when a job is released to the next cell, freeing a card: $D_i(t) = R_{i+1}(t)$. The policy is $\mathcal{B}_1 = \{x : x_1 + x_2^+ < c_1 + c_2\}$ and $\mathcal{B}_2 = \{x : x_1 > 0,$

$x_2 < c_2\}$.

Fixed Buffer

Under a finite buffer policy (see, e.g., Conway et al. 1988) the system operates as two tandem, finite capacity $\cdot/M/1/c_i$ queues. The buffer size is c_1 between stations and c_2 for FG; the policy is $\mathcal{B}_1 = \{x : x_1 < c_1\}$ and $\mathcal{B}_2 = \{x : x_1 > 0,\ x_2 < c_2\}$. This policy is also known as *fixed buffer* to distinguish it from a kanban policy where buffers are dynamically shifted as cards move in a cell, or *local control* because control of a station depends only on the number of jobs immediately downstream; i.e., station 1 is independent of the FG inventory.

CONWIP

The constant work-in-process (CONWIP) policy can be viewed as a kanban system with a single kanban cell (Muckstadt and Tayur 1991). For make-to-order or unlimited demand systems, CONWIP keeps the number of unfinished jobs in the system constant; for a make-to-stock system, the analogous policy is to keep WIP plus FG inventory constant. This policy is a special case of base stock with $c_1 = 0$.

3 Optimal Controls

It seems impossible to find a general solution to this control problem. For given parameter values λ, $\overline{\mu}_1$, $\overline{\mu}_2$, h and b, an optimal policy can be found numerically using dynamic programming; this is done in Section 5. One can also analyze a proposed policy π to determine conditions on the parameters under which it is optimal. The method used here to prove optimality is to establish bounds on $V^\pi(x + d_i) - V^\pi(x)$ using stochastic coupling arguments (see Hajek 1984 for another example of this technique). In special cases, these bounds can be used to verify that a policy is optimal.

For a given Markov policy π, let $\Delta_i(x) = V^\pi(x+e_i) - V^\pi(x)$ and $\Delta_{12}(x) = V^\pi(x-e_1+e_2) - V^\pi(x)$. The optimality condition (5) can be written

$$\begin{aligned} \Delta_1(x) &\leq 0 \text{ iff } x \in \mathcal{B}_1 &(13)\\ \Delta_{12}(x) &\leq 0,\ x \in \mathcal{B}_2 &(14)\\ \Delta_{12}(x) &\geq 0,\ x \in \mathcal{I}_2 \text{ and } x_1 > 0. &(15)\end{aligned}$$

Note that (14) and (15) only apply to points with $x_1 > 0$. We will write μ_i for $\overline{\mu}_i$ and μ_0 for λ when convenient. The subsections below make use of the following results. A cost of $c(x) = 1$ applied indefinitely yields a discounted cost of $1/\alpha$. Generalizing (5) to arbitrary policies gives

$$V^\pi(x) = \frac{1}{\Lambda(x)+\alpha}\left[c(x) + \sum_i \mu_i V^\pi(x+d_i)\right], \tag{16}$$

where $\Lambda(x)$ is the transition rate out of state x and the sum is taken over transitions i that are active in state x.

Let $(X(t), Y(t))$ be a coupled Markov process with state space

$$\mathbf{C} = \{(x,y) : x, y \in \mathbf{X} \text{ and } y = x, x+e_1, x+e_2, \text{ or } x-e_1+e_2\}, \tag{17}$$

where $X(t)$ and $Y(t)$ each have the same marginal distribution as the process of Section 1 under policy π with initial states $X(0)$ and $Y(0)$, and they share the same Poisson point processes of potential transitions. For example, in state $(x, x+e_1)$ the process transitions at rate μ_1 to $(x+e_1, x+e_1)$ if $x \in \mathcal{B}_1$ and $x+e_1 \in \mathcal{I}_1$, at rate μ_2 to $(x, x+e_2)$ if $x \in \mathcal{I}_2$ and $x+e_1 \in \mathcal{B}_2$, at rate μ_1 to $(x+e_1, x+2e_1)$ if $x, x+e_1 \in \mathcal{B}_1$, at rate μ_2 to $(x-e_1+e_2, x+e_2)$ if $x, x+e_1 \in \mathcal{B}_2$, and at rate λ to $(x-e_2, x+e_1-e_2)$.

Only the first two transitions change the relative position $Y(t) - X(t)$. We say that the process merges at the first time at which $Y(t) = X(t)$; then $Y(s) = X(s)$ for all $s \geq t$. We consider only policies that have the monotonicity properties of Fig. 3 (all optimal policies have these properties), thus limiting $Y(t) - X(t)$ to the values in $\mathbf{C}$. For example, if $x \in \mathcal{B}_1$ then $x - e_1 + e_2 \in \mathcal{B}_1$, and transitioning from $(x, x - e_1 + e_2)$ to $(x + e_1, x - e_1 + e_2)$ is impossible.

The usefulness of the coupled process lies in the fact that $\Delta_i(x)$ is the total cost resulting from a cost rate $c(Y(t)) - c(X(t))$ at time t, where $X(0) = x$ and $Y(0) = x + e_i$, and similarly for $\Delta_{12}(x)$, except that $Y(0) = x - e_1 + e_2$. The possible values of the cost rate are 0, 1, h, $h - 1$, $-b$, and $-b - 1$, corresponding to the values of $y - x$ for $(x, y) \in \mathbf{C}$.

3.1 No Inventory

Perhaps the simplest policy is to never hold inventory, releasing a job into station 1 only when there are backorders and station 2 is starved: $\mathcal{B}_1 = \{x : x_1 = 0,\ x_2 < 0\}$. Assume that station 2 is busy in states $(1,-1), (1,-2), (1,-3)\ldots$ so that the set of recurrent states is $\{(0,0), (0,-1), (0,-2), \ldots; (1,-1), (1,-2), (1,-3), \ldots\}$. For completeness, assume that the optimal control is used for station 2 in other, transient states. The task of checking optimality is much easier if only the recurrent states are checked. Although this condition is not sufficient for general Markov chains, it is for the chain defined here, as the following lemma shows.

Lemma 1 *For the no-inventory policy, (13) is implied by*

$$\Delta_1(0, x_2) \leq 0,\ x_2 < 0 \tag{18}$$

$$\Delta_1(1, x_2) \geq 0,\ x_2 < 0 \tag{19}$$

$$\Delta_1(0,0) \;\geq\; 0. \tag{20}$$

Proof. We must show that $\Delta_1(x) \geq 0$ when $x_1 > 1$ or $x_2 > 0$ (all cost functions are for the no-inventory policy). First, establish the condition for the states $\{(0,x_2) : x_2 > 0\}$ using induction on x_2. To evaluate $\Delta_1(0,x_2)$, we must consider the states $(1,x_2)$. Since station 2 uses the optimal control in these states, monotonicity holds: station 2 is busy up to some x_2 and idle beyond. For states $(1,x_2) \in \mathcal{B}_2$, we will establish that $\Delta_1(0,x_2)$ is increasing. Initially $\Delta_1(0,0) \geq 0 \geq \Delta_1(0,-1)$. Assume that $\Delta_1(0,x_2) \geq \Delta_1(0,x_2-1)$ for some $x_2 > 0$. Since $(0,x_2+1) \in \mathcal{I}_1 \cap \mathcal{I}_2$ and $(1,x_2+1) \in \mathcal{I}_1 \cap \mathcal{B}_2$, there are two transitions for the coupled process corresponding to $\Delta_1(0,x_2+1)$, and (16) yields

$$\Delta_1(0,x_2+1) = \frac{1}{\lambda+\mu_2+\alpha}\left[1 + \lambda\Delta_1(0,x_2) + \mu_2\Delta_2(0,x_2+1)\right]. \tag{21}$$

Similarly, using the transitions from $(x, x+e_2)$ gives

$$\Delta_2(0,x_2+1) = \frac{1}{\lambda+\alpha}\left[h + \lambda\Delta_2(0,x_2)\right]. \tag{22}$$

Since the cost rate for the coupled process is never more than h, $\Delta_2(0,x_2) \leq h/\alpha$; eliminating h in (22) gives $\Delta_2(0,x_2+1) \geq \Delta_2(0,x_2)$. Using this fact and the inductive hypothesis in (21),

$$\begin{aligned}
\Delta_1(0,x_2+1) \;&\geq\; \frac{1}{\lambda+\mu_2+\alpha}\left[1 + \lambda\Delta_1(0,x_2-1) + \mu_2\Delta_2(0,x_2)\right] && (23)\\
&=\; \Delta_1(0,x_2). && (24)
\end{aligned}$$

Therefore $\Delta_1(0,x_2) \geq \Delta_1(0,0) \geq 0$ for $(1,x_2) \in \mathcal{B}_2$. For $(1,x_2) \in \mathcal{I}_2$, the last term in (21) is omitted and a simple induction argument shows that $\Delta_1(0,x_2+1)$ remains nonnegative.

Now consider $\Delta_1(x)$ for states with $x_1 > 0$. The corresponding coupled process has a cost rate of one until the first time t such that $X(t) = y$ for some y such that $y_1 = 1$ and $y_2 < 0$, or $y_1 = 0$ and $y_2 > 0$. Thereafter the cost rate is the same as for $\Delta_1(y)$. But $\Delta_1(y) \geq 0$ for such y; hence, $\Delta_1(x) \geq 0$. □

Using rather crude stochastic coupling bounds, the following parameter ranges are obtained from (14), (15), and (18-20).

Theorem 1 *The following conditions are sufficient for the no-inventory policy to be optimal:*

$$b\mu_2 \geq \mu_1 + \alpha, \tag{25}$$

$$1 + \mu_2 \min\{\phi, 0\} + \frac{\mu_2}{\mu_1 + \mu_2 + \alpha}\left[1 - \frac{\mu_2(b+1)}{\alpha}\right] \leq 0, \tag{26}$$

$$\text{and } 1 + \frac{\lambda}{\mu_1 + \mu_2 + \alpha}\left[1 - \frac{\mu_2(b+1)}{\alpha}\right] + \frac{\mu_2}{\lambda + \alpha}\left[h - \frac{\lambda(b+1)}{\alpha}\right] \geq 0, \tag{27}$$

$$\text{where } \phi = \left(\frac{\mu_2}{\mu_2 + \alpha}\right)\frac{h}{\lambda + \alpha} + \left(\frac{\mu_1 + \mu_2}{\mu_1 + \mu_2 + \alpha}\right)\left(\frac{\mu_2}{\mu_1 + \mu_2} - \frac{\lambda}{\lambda + \alpha}\right)\frac{b+1}{\alpha}. \tag{28}$$

The proof of Theorem 1, which is a lengthy application of the coupled process, is given in the Appendix. Conditions (25-27) only hold for very large α, on the order of μ_2; one example is $h = 2$, $b = 4$, $\mu_1/\alpha = 2$, $\mu_2/\alpha = 1$, and $\lambda/\alpha = 1/2$. This result is reasonable because a policy of not holding inventory is very shortsighted. It can be optimal only if the time horizon $1/\alpha$ is sufficiently short.

3.2 No FG Inventory

Now consider a policy that consists of the optimal control of station 1 and operating station 2 only when there are backorders: $\mathcal{B}_2 = \{x : x_1 > 0, x_2 < 0\}$. A lemma again allows us to omit the transient states $\{x : x_2 > 0\}$ from the optimality check. Here the assumption that $h > 1$ is critical; otherwise all WIP would be converted to FG.

Lemma 2 *For the no-FG inventory policy, $\Delta_{12}(x_1, 0) \geq 0$ implies (15).*

Proof. Again using the coupled process, $\Delta_{12}(x)$ for $x_2 > 0$ corresponds to the cost rate $h - 1$ until $X_2(t) = 0$, then the same costs as $\Delta_{12}(y_1, 0)$ for some y_1. Both of these costs are nonnegative. □

Theorem 2 *The following conditions are sufficient for the no-FG inventory policy to be optimal:*

$$\mu_2 \geq \lambda \left(\frac{\lambda + \alpha}{\alpha}\right) \left(\frac{b+1}{h-1}\right) - \lambda - \alpha, \tag{29}$$

$$h - 1 + \frac{\lambda}{\lambda + \mu_2 + \alpha}(-b - 1 + \lambda V_5 + \mu_2 V_3) \geq 0, \tag{30}$$

$$\text{and } \ h - 1 + \lambda V_2 \geq 0, \tag{31}$$

$$\begin{aligned}
\text{where } V_3 &= \theta_1[\lambda + \mu_2 + \alpha + \lambda(1 + \lambda V_5)] && (32)\\
V_5 &= -(b+1)/\alpha && (33)\\
\theta_1 &= \frac{1}{(\lambda + \alpha)(\lambda + \mu_2 + \alpha) - \lambda\mu_2} && (34)\\
V_2 &= \theta_2 \left[-b - 1 + \lambda V_5 + \frac{\mu_2}{\lambda + \alpha}\left(1 + \frac{\lambda(1 + \lambda V_5)}{\lambda + \mu_2 + \alpha}\right)\right] && (35)\\
\theta_2 &= \frac{(\lambda + \alpha)(\lambda + \mu_2 + \alpha)}{(\lambda + \alpha)(\lambda + \mu_2 + \alpha)^2 - \lambda\mu_2^2}. && (36)
\end{aligned}$$

In particular, it is optimal for sufficiently large μ_2.

The proof in the Appendix suggests that other conditions for optimality could be obtained if desired.

3.3 Non-Idling at Station 2

In some sense the opposite of the previous policy is to never idle at station 2: $\mathcal{B}_2 = \{x : x_1 > 0\}$ and optimal control of station 1. Reversing the cost structure so that it is more costly to hold WIP than FG makes this policy optimal.

Theorem 3 *If $h < 1$ then non-idling at station 2 is optimal.*

Proof. Suppose a policy π includes idling at station 2 in some state x with $x_1 > 0$. For this initial state x, construct a policy π' that is identical to π except that the start time of the next job at station 2 is moved up to zero. Their cost rates differ by $h - 1$ or $-b - 1$, both nonpositive, for the time interval between the completion of the next job at station 2 under π' and under π. Therefore $V^{\pi'}(x) < V^{\pi}(x)$ and only policies that are non-idling at station 2 are optimal. □

Theorem 3 can be strengthened to the case $h = 1$ using a more elaborate proof. A situation where $h \leq 1$ might occur when the benefits of just-in-time manufacturing are incorporated as additional WIP holding costs.

When $h > 1$ the decision of whether to operate station 2 depends on the likelihood of incurring FG holding costs as a result. If the optimal control for station 1 prevents FG inventory from being held, then non-idling is optimal at station 2.

Theorem 4 *If $\mathcal{B}_1 \cap \{x : x_2 \geq 0\} = \emptyset$ then it is optimal not to idle at station 2 in all states that are recurrent under some station 2 control.*

Proof. Let $\overline{\mathcal{B}}_1 = \{x : x \in \mathcal{B}_1 \text{ or } x - d_1 \in \mathcal{B}_1\}$. Recall from Section 1 that $\mathcal{B}_1$, and also $\overline{\mathcal{B}}_1$, consists of all states below a switching curve $s_1(x_1)$ with slope $s_1'(x_1) \leq -1$. Thus, regardless of the station 2 control the system cannot leave $\overline{\mathcal{B}}_1$. Other states cannot be recurrent because the transition d_1 cannot occur outside of $\overline{\mathcal{B}}_1$. Now, for x recurrent with $x_1 > 0$, we must have $x \in \overline{\mathcal{B}}_1$ and $x_2 < 0$. Hence, from the initial state x, $X(t) \in \overline{\mathcal{B}}_1$ for all t and the coupled process cost rate for $\Delta_{12}(x)$ is either $-b-1$ or $-b$. This implies that $\Delta_{12}(x) \leq 0$ and non-idling is optimal at station 2. □

The above proof also generalizes the proof of (14) for the no-inventory policy of Theorem 1.

3.4 Base Stock Policies are Never Optimal

Despite their popularity in inventory systems, base stock policies are never optimal for this problem because they can accumulate large amounts of WIP that, due to the capacity constraint, will remain in the system for long periods of time.

Theorem 5 *The base stock policy of Section 2 is not optimal.*

Proof. Consider a base stock policy with $\mathcal{B}_1 = \{x : x_1 + x_2 < c_1 + c_2\}$ and $\mathcal{B}_2 = \{x : x_1 > 0,\ x_2 < c_2\}$. For the coupled process associated with $\Delta_1(x)$, let T_1 be the time of first departure from a cost rate of one, T_m be the merge time, $p(x) = \Pr\{T_1 < T_m\}$, V_1 be the discounted cost until T_1, and V_2 be the discounted cost from T_1 to T_m given that $T_1 < T_m$. Then

$$\Delta_1(x) = V_1 + p(x)V_2. \tag{37}$$

We will show that, for $x = (x_1, c_1 + c_2 - x_1 - 1)$ and $x_1 \to \infty$, $p(x) \to 0$. Since $V_1 > 0$ and $V_2 \leq h/\alpha$, this implies that $\Delta_1(x) \geq 0$ for some $x = (x_1, c_1 + c_2 - x_1 - 1)$ and the base stock policy is not optimal.

From the initial state x, the event $\{T_1 < T_m\}$ requires $X_1(t) = 0$ for some $t < T_m$; hence, no merge can occur during the first x_1 potential transitions of X. Let N_0,

N_1, and N_2 be the number of the first x_1 transitions that are of type λ, μ_1, and μ_2, respectively, and T be the time of the x_1-th transition. Then

$$\begin{aligned} p(x) &\leq \Pr\{T_m > T\} && (38) \\ &\leq \Pr\{N_1 - N_0 < 1\}, && (39) \end{aligned}$$

since $N_1 - N_0 \geq 1$ implies that the first occurance of $X_1(t) + X_2(t) = c_1 + c_2$ was at some time $t \leq T$ and the process merged at that time. But (N_0, N_1, N_2) have a multinomial distribution with x_1 trials and probabilities $(\lambda/\Lambda, \mu_1/\Lambda, \mu_2/\Lambda)$, so that $\mu_1 > \lambda$ implies that $\Pr\{N_1 - N_0 < 1\} \to 0$ as $x_1 \to \infty$. □

4 An Equivalent Make-to-Order System

The production/inventory system of Section 1 is not a tandem queue in the usual sense because backorders ($x_2 < 0$) are allowed; instead, it has been viewed as a queueing network with assembly where demands enter a queue and are joined with FG (Mitra and Mitrani 1991). However, if total inventory is bounded, we can transform this system into an equivalent make-to-order system. Consider only policies π and initial states $X(0) = (c_1, c_2)$ for which $X_1(t) + X_2(t) \leq c_1 + c_2$, or equivalently, $S_1(t) \leq D(t)$, with probability one. Define

$$\begin{aligned} Z_1(t) &= c_1 + c_2 - X_1(t) - X_2(t) && (40) \\ Z_2(t) &= X_1(t). && (41) \end{aligned}$$

Then Z is a make-to-order system (a tandem queue) with infinite first buffer; rates λ, μ_1, and μ_2; and cost function $c^Z(z) = c^X(z_2, c_1 + c_2 - z_1 - z_2) = z_2 + h(c_1 + c_2 - z_1 - z_2)^+ + b(c_1 + c_2 - z_1 - z_2)^-$. All statistics of X can be recovered from Z by solving (40) and (41) for X. The distinguishing feature of a make-to-stock system is seen to be its concave (as opposed to linear) cost function, not its dynamics. Such a transformation can be made for n-stage systems as well.

Because of this equivalence, any method that obtains the state probability distribution for tandem queues can be used to evaluate a make-to-stock system under the corresponding control policy. The fixed buffer mechanism provides a good example. A fixed buffer of size c_1 (with infinite first buffer) for Z gives the policy $\mathcal{B}_1 = \{x : x_1 + x_2 < c_1 + c_2,\ x_1 < c_1\}$ and $\mathcal{B}_2 = \{x : x_1 > 0\}$ for X, where c_2 is arbitrary. This control mechanism, motivated by the linear cost function of make-to-order systems, will not always be appropriate for the concave cost function. However, when it is reasonable, the following method could be used to obtain a nearly optimal policy of this form. For a given c_1, generate a steady-state distribution using one of the approximations noted in Smith and Daskalaki (1988). Then, given c_2, compute the appropriate cost measure. Use an optimization scheme to find the best c_1 and c_2.

5 Dynamic Programming Computational Results

Dynamic programming value iteration was used on a truncated state space to compute the optimal policy for several cases. For undiscounted problems, the average cost per unit time g/Λ is reported; in the discounted case, the cost $V(0,0)$ is reported. Up to 2000 iterations were required to achieve four digit accuracy. Larger and larger state spaces were tested until the results were insensitive to increasing the state space. The largest state space required was 21 by 43. To avoid solving large linear systems, value iteration was also used to evaluate candidate policies. A coordinate search algorithm was employed to find the best parameters (c_1, c_2) for a given type of policy. The algorithm assumes convexity of V^π; to check this assumption different initial values

Table 1: Gain per Unit Time Under Various Policies ($\lambda = 1$, $h = 2$, $b = 4$, $\alpha = 0$)

Policy	Case 1 $\mu_1 = \mu_2 = 1.2$	Case 2 $\mu_1 = 2,\ \mu_2 = 1.2$	Case 3 $\mu_1 = 1.2,\ \mu_2 = 2$
Optimal	21.50	14.88	11.48
Best Base Stock	21.57	15.9	11.56
Best Kanban	22.1	15.3	11.63
Best Fixed Buffer	23.7	16.4	11.83
Revised Base Stock	21.54	15.2	11.56

Table 2: Suboptimality and Stock Levels for Various Policies ($\lambda = 1$)

		Base Stock		Kanban		Fixed Buffer	
Case	$(\mu_1, \mu_2, h, b, \alpha)$	% Subopt	(c_1, c_2)	% Subopt	(c_1, c_2)	% Subopt	(c_1, c_2)
1	1.2,1.2,2,4,0.00	0.3	4,8	3.0	6,8	10	12,7
2	2.0,1.2,2,4,0.00	7	1,6	3	1,6	10	5,6
4	1.2,1.2,2,4,0.10	1.2	1,3	0.8	2,3	2.2	5,3
5	2.0,2.0,2,4,0.00	0.9	1,2	5.5	1,2	17	3,1
6	2.0,1.2,1,1,0.00	24	1,3	6	1,4	15	4,4

of (c_1, c_2) were tried and gave the same results.

Three undiscounted cases are presented in Table 1. Case 1 is a balanced system with a utilization of 5/6. In case 2 station 1 is faster, while in case 3 station 2 is faster. As is known for a variety of manufacturing systems, it is better to have the faster machine downstream so that the bottleneck is upstream (case 3). Among the suboptimal policies, base stock performs very well for cases 1 and 3. When the utilization of the upstream machine is low, as in case 2, stockpiling WIP when there are many backorders is unnecessary and the base stock policy does not perform as well as kanban. These results are also explained by Figs. 5, 6, and 7, showing the optimal busy regions. Case 3 has the 45-degree line characteristic of base stock policies

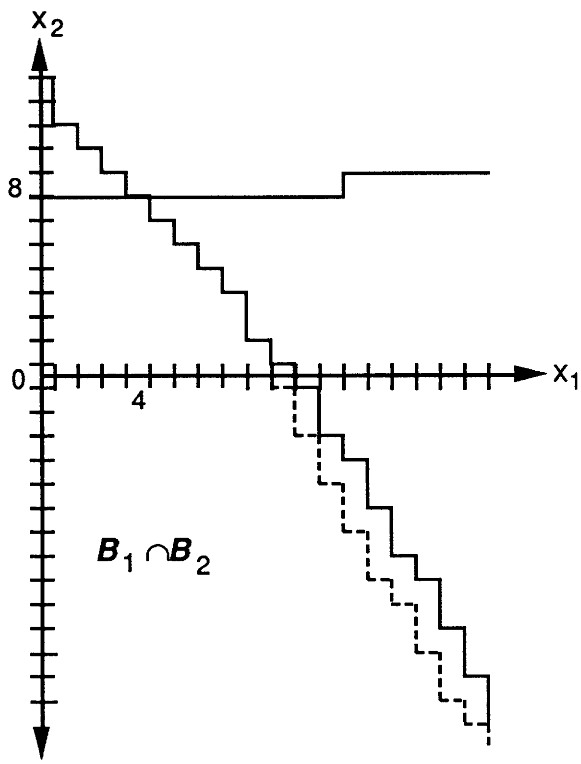

Figure 5: Optimal Policy for Case 1 (dashed line is revised base stock)

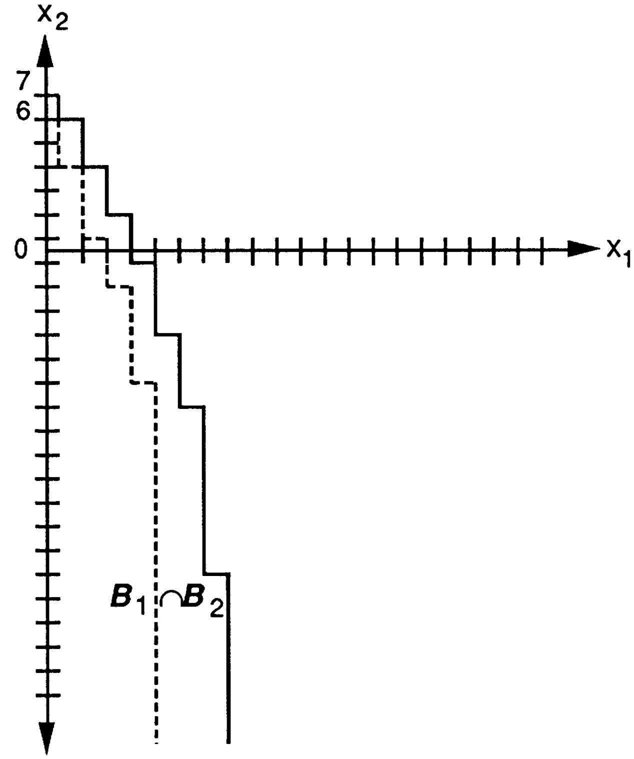

Figure 6: Optimal Policy for Case 2 (dashed line is case 6)

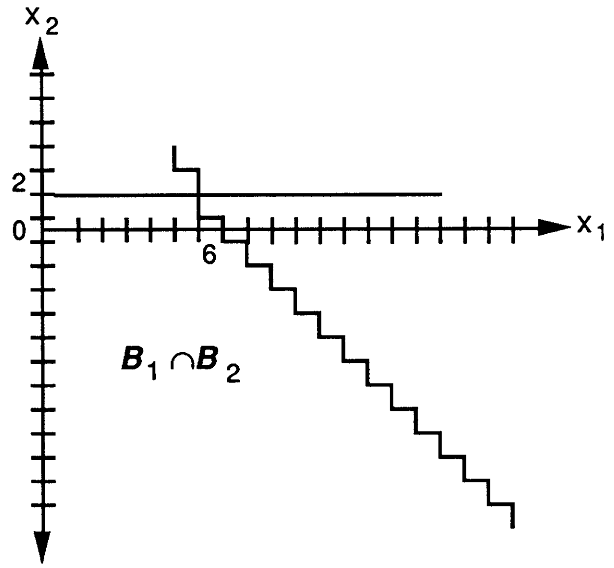

Figure 7: Optimal Policy for Case 3

(although, by Theorem 5, this line must turn downward for large x_1), case 1 is nearly base stock, and case 2 is quite different.

Table 2 compares the policies in greater detail for additional test cases. The parameters c_1 and c_2 are the hedging point; i.e., the target levels of WIP and FG, respectively. Case 4 illustrates that much less stock is held and kanban is preferable when discounting is present (compare case 1). Case 5 shows that little stock is held when utilizations are low. In case 6, the combination of a faster downstream station and relatively high WIP holding costs creates a situation where significant FG but little WIP is held. The optimal policy for case 6, shown in Fig. 6, allows up to six units of FG to be held but has an even steeper switching curve than case 2.

The good performance of base stock policies suggests a way of quickly generating a nearly optimal policy: search over base stock policies to find the best one, evaluating V^{π} for each policy. Then revise the station 1 switching curve for this policy using V^{π} by setting $x \in \mathcal{B}_1$ if $\Delta_1(x) \leq 0$. Performance of this "revised base stock" policy is

Table 3: Configurations of the Coupled Process in (18)

Configuration	States (X,Y)	Cost Rate	Departure Rate Λ
1	$(x, x+e_1)$	1	$\mu_1+\mu_2$
2	$(x, x+e_2)$	$-b$	$0 \leq \Lambda \leq \mu_1$
3	$(x, x-e_1+e_2)$	$-b-1$	$\mu_2 \leq \Lambda \leq \mu_1+\mu_2$
4	(x, x)	0	0

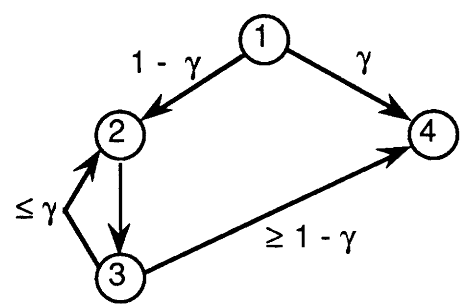

Figure 8: Transition Diagram for the Coupled Process in (18); $\gamma = \mu_1/(\mu_1+\mu_2)$.

included in Table 1. As shown in Fig. 6, for case 1 this policy has nearly the same busy region as the optimal policy. Although we have used a slow, iterative algorithm to compute V^π in this study, a rapid solution should be possible because of the sparse structure of the linear system (16).

Appendix

Proof of Theorem 1.

In light of Lemma 1, it suffices to show (18-20) and $\Delta_{12}(1,x_2) \leq 0$, $x_2 < 0$.

(18). Consider the coupled process associated with $\Delta_1(0,x_2)$, $x_2 < 0$. Partition its

Table 4: Configurations of the Coupled Process in (19)

Configuration	States (X,Y)	Cost Rate	Departure Rate Λ
1	$(x, x+e_1)$, $x_1 = 1$	1	μ_2
2	$(x, x+e_1)$, $x_1 = 0$	1	$\mu_2 \leq \Lambda \leq \mu_1 + \mu_2$
3	$((0,0), e_2)$	h	λ
4	$(x, x+e_2)$, $x_2 < 0$	$-b$	–

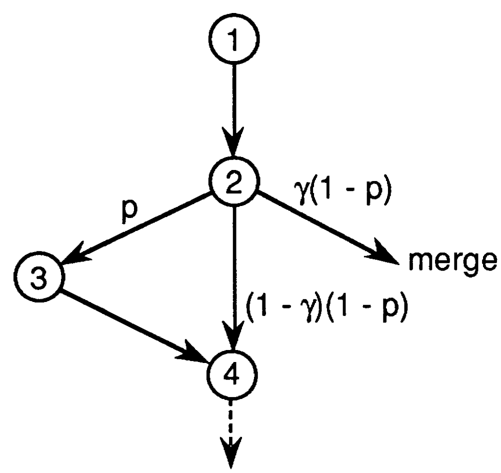

Figure 9: Transition Diagram for the Coupled Process in (19); $\gamma = \mu_1/(\mu_1 + \mu_2)$.

Table 5: Configurations of the Coupled Process in (20)

Configuration	States (X,Y)	Cost Rate	Departure Rate Λ
1	$((0,0), e_1)$	1	$\lambda + \mu_2$
2	$((0,0), e_2)$	h	λ
3	$(x, x+e_1)$, $x_2 < 0$	1	$\mu_1 + \mu_2$
4	$(x, x+e_2)$, $x_2 < 0$	$-b$	–

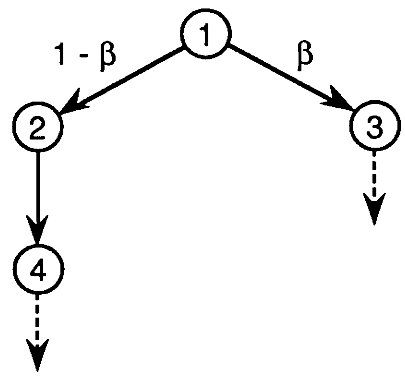

Figure 10: Transition Diagram for the Coupled Process in (20); $\beta = \lambda/(\lambda + \mu_2)$.

possible states into the configurations listed in Table 3. It moves through these configurations according to Fig. 8, where the quantity next to each arc is the probability of moving along that arc. Let T_i be the time of first departure from configuration i, T_m be the merge time (merge occurs upon entering configuration 4), and $V(s,t)$ be the cost incurred by the coupled process in the period $(s,t]$, given that the process has not merged by s. Then

$$\begin{aligned}
V(T_1,T_3) &= E\left[-b\int_0^{T_2-T_1} e^{-\alpha t}\mathrm{d}t \;-\; (b+1)e^{-\alpha(T_2-T_1)}\int_0^{T_3-T_2} e^{-\alpha t}\mathrm{d}t\right] \\
&\leq E\left[-\frac{b}{\alpha}\left(1-e^{-\alpha(T_2-T_1)}\right) - \frac{b+1}{\mu_1+\mu_2+\alpha}e^{-\alpha(T_2-T_1)}\right] \\
&\leq \max_{0\leq t\leq\infty}\left[-\frac{b}{\alpha}\left(1-e^{-\alpha t}\right) - \frac{b+1}{\mu_1+\mu_2+\alpha}e^{-\alpha t}\right] \\
&= -\min\left\{\frac{b}{\alpha}, \frac{b+1}{\mu_1+\mu_2+\alpha}\right\}. \qquad \text{(A.1)}
\end{aligned}$$

The first integral was evaluated exactly; the second was bounded using the maximum departure rate. From (16) and Fig. 8,

$$\Delta_1(0,x_2) \leq \frac{1}{\mu_1+\mu_2+\alpha}\left[1-\mu_2\min\left\{\frac{b}{\alpha}, \frac{b+1}{\mu_1+\mu_2+\alpha}\right\}\right], \qquad \text{(A.2)}$$

where we have omitted $V(T_3,T_m) < 0$. The right side is nonpositive when (25) holds, i.e., (25) implies (18).

(19). For $\Delta_1(1, x_2)$, $x_2 < 0$, the relevant configurations are listed in Table 4, with the transitions in Fig. 9. Here p is the probability that $X(T_2^-) = (0,0)$ is the last state visited in configuration 2. A crude bound after entering configuration 4 is $V(T_3, T_m) \geq -(b+1)/\alpha$ (define $T_3 = T_2$ if configuration 3 is not visited). Denote the discount factor while in configuration i by $\alpha_i = E\{\exp[-\alpha(T_i - T_{i-1})]\}$. Using the maximum and minimum rates from Table 4 where applicable, and applying (16) repeatedly gives

$$\begin{aligned}\Delta_1(1,x_2) &= V(0,T_1) + \alpha_1 \left\{ V(T_1,T_2) + \alpha_2 \left[pV(T_2,T_3) + p\alpha_3 V(T_3,T_m) \right.\right.\\ &\quad \left.\left. + \frac{\mu_2}{\mu_1+\mu_2}(1-p)V(T_3,T_m) \right] \right\} \\ &\geq \frac{1}{\mu_2+\alpha}\left\{ 1 + \frac{\mu_2}{\mu_1+\mu_2+\alpha} + \frac{\mu_2^2 ph}{(\mu_2+\alpha)(\lambda+\alpha)} \right. \\ &\quad \left. - \left(\frac{\mu_2(\mu_1+\mu_2)}{\mu_1+\mu_2+\alpha}\right)\left[(1-p)\left(\frac{\mu_2}{\mu_1+\mu_2}\right)\frac{b+1}{\alpha} + p\left(\frac{\lambda}{\lambda+\alpha}\right)\frac{b+1}{\alpha}\right]\right\}. \end{aligned} \tag{A.3}$$

Taking the minimum over $0 \leq p \leq 1$,

$$\Delta_1(1,x_2) \geq \frac{1}{\mu_2+\alpha}\left\{1 + \mu_2 \min\{\phi, 0\} + \frac{\mu_2}{\mu_1+\mu_2+\alpha}\left[1 - \frac{\mu_2(b+1)}{\alpha}\right]\right\}, \tag{A.4}$$

$$\text{where } \phi = \left(\frac{\mu_2}{\mu_2+\alpha}\right)\frac{h}{\lambda+\alpha} + \left(\frac{\mu_1+\mu_2}{\mu_1+\mu_2+\alpha}\right)\left(\frac{\mu_2}{\mu_1+\mu_2} - \frac{\lambda}{\lambda+\alpha}\right)\frac{b+1}{\alpha}.$$

The right side of (A.4) is nonnegative when (26) holds.

(20). For $\Delta_1(0,0)$, a partial list of configurations and their transitions are given in Table 5 and Fig. 10. The first transition gives

$$\Delta_1(0,0) = \frac{1}{\lambda+\mu_2+\alpha}\left[1 + \lambda\Delta_1(0,-1) + \mu_2\Delta_2(0,0)\right]. \tag{A.5}$$

Table 6: Configurations of the Coupled Process in (15)

Configuration	States (X,Y)	Cost Rate	Departure Rate Λ
1A	$(x, x-e_1+e_2)$, $x_2=0$	$h-1$	λ
1B	$(x, x+e_2)$, $x_2=0$	h	λ
2A	$(x, x-e_1+e_2)$, $x_2=-1$	$-b-1$	$\lambda+\mu_2$
2B	$(x, x+e_2)$, $x_2=-1$	$-b$	$\lambda+\mu_2$
3	$(x, x+e_1)$, $x_2=0$	1	$\lambda \le \Lambda \le \lambda+\mu_1$
4	$(x, x+e_1)$, $x_2=-1$	1	$\lambda \le \Lambda \le \lambda+\mu_1+\mu_2$
5	(x,y), $x_2<-1$	–	–

Change (A.2) to a lower bound using $V(T_2, T_m) \ge -(b+1)/\alpha$, giving

$$\Delta_1(0,-1) \ge \frac{1}{\mu_1+\mu_2+\alpha}\left[1 - \frac{\mu_2(b+1)}{\alpha}\right]. \tag{A.6}$$

Expanding $\Delta_2(0,0)$ using the same lower bound, (A.5) can be written

$$\Delta_1(0,0) \ge \frac{1}{\lambda+\mu_2+\alpha}\left\{1 + \frac{\lambda}{\mu_1+\mu_2+\alpha}\left[1-\frac{\mu_2(b+1)}{\alpha}\right] + \frac{\mu_2}{\lambda+\alpha}\left[h - \frac{\lambda(b+1)}{\alpha}\right]\right\}. \tag{A.7}$$

The right side is nonnegative when (27) holds.

$\Delta_{12}(1,x_2)$, $x_2<0$. The cost rate of the corresponding coupled process is negative, $-b$ or $-b-1$, until merging. Hence, $\Delta_{12}(1,x_2) \le 0$. □

Proof of Theorem 2.

By assumption, (13) holds. The proof of (14) is identical to that of Theorem 3 except that $x_2 < 0$ and the cost rates differ by $-b-1<0$, since x_2 can only decrease until the next job is processed at station 2. In light of Lemma 2, it remains to show (15) for $x_2 = 0$. Consider the configurations in Table 6 and the transitions in Fig. 11 for the coupled process corresponding to $\Delta_{12}(x_1, 0)$. Here the cost rate

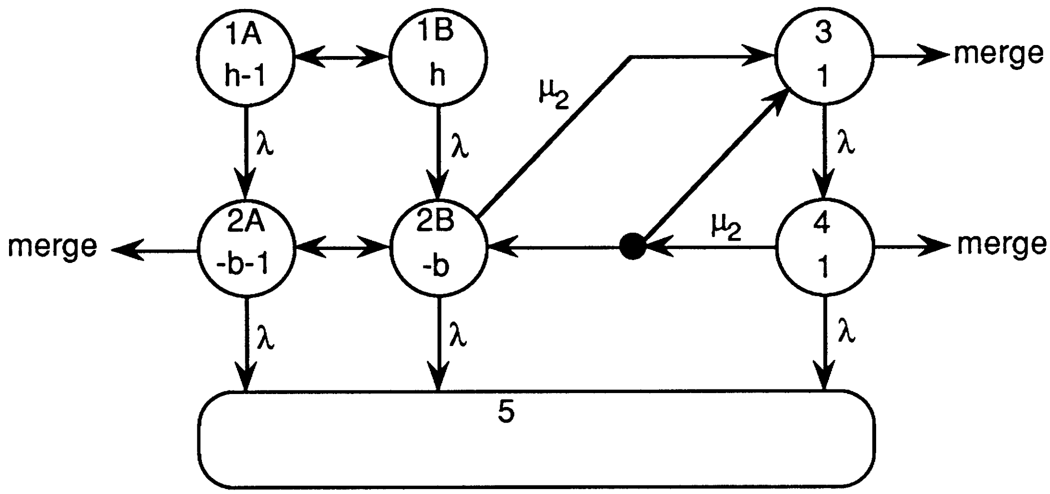

Figure 11: Transition Diagram for the Coupled Process in (15)

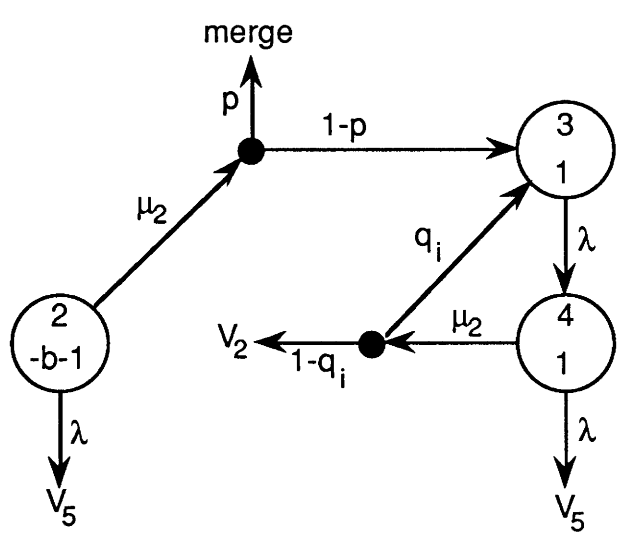

Figure 12: Lower Bound Process

appears in each configuration and the transition rate appears next to each arc that has a constant rate (recall that the configurations are collections of states, so that the configuration does not evolve as a Markov process). The definition of these configurations is motivated by the fact that, as μ_2 increases, the probability of entering configuration 5 ($x_2 < -1$) decreases and more time is spent in configuration 3 with a positive cost. Adopt the notation $\mathcal{C}_i = \{(x,y) : (x,y) \in \text{configuration } i\}$, $V(x,y) = V^\pi(y) - V^\pi(x)$ (the cost incurred by the coupled process under the no-FG policy π), and $V_i = \min_{(x,y)\in\mathcal{C}_i} V(x,y)$. Then

$$\Delta_{12}(x_1,0) = V((x_1,0),(x_1-1,1)) \geq V_{1A} \geq \frac{1}{\lambda+\alpha}(h-1+\lambda V_2). \tag{A.8}$$

To bound V_2, we will approximate Fig. 11 with the process of Fig. 12. For a given state $(x,y) \in \mathcal{C}_2$, let p be the probability of merging before returning to $\mathcal{C}_2$, given that the system leaves $\mathcal{C}_2$ by a μ_2 transition, and let q_i be the probability of returning to configuration 3 upon leaving configuration 4 (by a μ_2 transition) for the ith time. We claim that, for this p and q_i, the approximate process is a lower bound in the sense that $V(x,y) \geq V_2'$, where V_2' is the value of the approximate process in state 2. Applying this bound for all states in $\mathcal{C}_2$ gives

$$V_2 \geq V_2'. \tag{A.9}$$

To establish that the approximate process is a lower bound, make the following sequence of changes to Fig. 11, each of which decreases (or does not change) V_2. Eliminate configuration 1, replace the cost upon entering configurations 2B and 5 with their lower bounds V_2 and V_5, change the cost in 2B to $-b-1$, move the merges after configurations 3 and 4 to after configuration 2B (with an equivalent probability of merging; the only effect is to reduce the time spent in 3 and 4), and combine 2A and 2B into 2 (the equivalent probability of merging is p, defined above).

Analyzing Fig. 12,

$$V_2' = \frac{1}{\lambda + \mu_2 + \alpha}[-b - 1 + \lambda V_5 + (1-p)\mu_2 V_3']. \tag{A.10}$$

To establish bounds, use $V_5 \geq -(b+1)/\alpha$ and consider the values $p = 0$ and 1, one of which will be a worst case. If $p = 1$,

$$\Delta_{12}(x_1, 0) \geq \frac{1}{\lambda + \alpha}\left(h - 1 - \frac{\lambda}{\lambda + \mu_2 + \alpha}\left[b + 1 + \frac{\lambda(b+1)}{\alpha}\right]\right). \tag{A.11}$$

The right side is nonnegative if (29) holds.

Now consider $p = 0$. Either $q_i = 0$ for all i or $q_i = 1$ for all i is a worst case (other than variation in the transition probabilities q_i, the system is Markov, so that the minimal cost is achieved by a constant q). If $q_i = 1$,

$$V_3' = \frac{1}{\lambda + \alpha}\left[1 + \frac{\lambda}{\lambda + \mu_2 + \alpha}(1 + \lambda V_5 + \mu_2 V_3')\right]. \tag{A.12}$$

Solving for V_3' and dropping the prime notation gives (32). Combining (A.8-10) and requiring the right side to be nonnegative gives (30). Hence, for the case $q_i = 1$, (30) implies (15). If $q_i = 0$, V_2 replaces V_3' in the right side of (A.12). In light of (A.9), we can replace V_2 with V_2', substitute into (A.10), and solve for V_2' to obtain the lower bound (35). Since (31) requires the right side of (A.8) to be nonnegative, it implies (15) for the case $q_i = 0$. □

References

Bertsekas, D.P. 1976. *Dynamic Programming and Stochastic Control.* Academic Press, New York.

Bielecki, T. and P.R. Kumar 1988. Optimality of Zero-Inventory Policies for Unreliable Manufacturing Systems. *Oper. Res.* **36**, 532-541.

Buzacott, J., S. Price, and J.G. Shanthikumar 1991. Service Levels in Multistage MRP and Base Stock Controlled Production Systems. Working paper.

Cheng, D.W. and D.D. Yao 1991. Tandem Queues with Generalized Blocking: A Unified Model and Comparison Results. Working paper, New York University, New York.

Conway, R., W. Maxwell, J.O. McClain, and L.J. Thomas 1988. The Role of Work-in-Process Inventory in Serial Production Lines. *Oper. Res.* **36**, 229-241.

Hajek, B. 1984. Optimal Control of Two Interacting Service Stations. *IEEE Trans. Auto. Control* **AC-29**, 491-499.

Hillier, F.S., R.W. Boling, and K.C. So 1986. Toward Characterizing the Optimal Allocation of Storage Space in a Production Line System with Variable Operating Times. Technical Report TR-76, Stanford University, Stanford.

Lee, Y.J. and P. Zipkin 1990. Tandem Queues with Planned Inventories. Working paper, Graduate School of Business, Columbia University, New York.

Lippman, S. 1975. Applying a New Device in the Optimization of Exponential Queueing Systems. *Oper. Res.* **23**, 687-710.

Lou, S.X.C. and G. van Ryzin 1989. Optimal Control Rules for Scheduling Job Shops. *Annals Oper. Res.* **17**, 233-248.

Mitra, D. and I. Mitrani 1990. Analysis of a Kanban Discipline for Cell Coordination in Production Lines, I. *Mgmt. Science* **36**, 1548-1566.

Mitra, D. and I. Mitrani 1991. Analysis of a Kanban Discipline for Cell Coordination in Production Lines, II: Stochastic Demands. *Oper. Res.* **39**, 807-823.

McClain, J.O. and D.R. Moodie 1991. A Comment on "Buffer Space Allocation in Automated Assembly Lines." *Oper. Res.* **39**, 857-860.

Muckstadt, J.A. and S.R. Tayur 1991. A Comparison of Alternative Kanban Control Mechanisms. Technical Report 962, School of Oper. Res. and Indust. Engr., Cornell University, Ithaca.

Perros, H.G. 1984. Queueing Networks with Blocking: A Bibliography. Performance Evaluation Review, *ACM Sigmetrics* **12**, 8-12.

Rosberg, Z., P. Varaiya, and J. Walrand 1982. Optimal Control of Service in Tandem Queues. *IEEE Trans. Auto. Control* **AC-27**, 600-610.

Schonberger, R.J. 1982. *Japanese Manufacturing Techniques.* Free Press, Detroit.

Smith, J.M. and S. Daskalaki 1988. Buffer Space Allocation in Automated Assembly Lines. *Oper. Res.* **36**, 343-358.

Sugimori, Y., K. Kusunoki, F. Cho, S. Uchikawa 1977. Toyota Production Systems and Kanban System-Materialization of Just-in-Time and Respect-for-Human System. *Int. J. Prod. Res.* **15**, 553-564.

van Ryzin, G., S.X.C. Lou, and S.B. Gershwin 1991. Production Control for a Tandem Two-Machine System. Laboratory for Manufacturing and Productivity Report LMP-91-001, M.I.T., Cambridge.

Veatch, M.H. and L.M. Wein 1991. Monotone Control of Queueing Networks. Working paper OR-257-91, Operations Research Center, M.I.T., Cambridge.

Weber, R. and S. Stidham 1987. Optimal Control of Service Rates in Networks of Queues. *Adv. Appl. Prob.* **19**, 202-218.