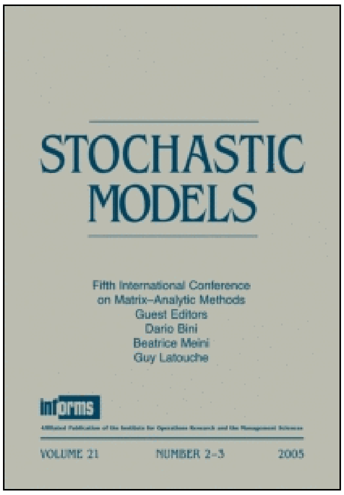

Stochastic Models
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597301

A lost-sales production/inventory model with two discrete production modes

Rein D. Nobel[a]; Mattijs van der Heeden[a]
[a] Department of Econometrics, Faculty of Economics and Econometrics, Vrije Universiteit, The Netherlands

To cite this Article Nobel, Rein D. and van der Heeden, Mattijs(2000) 'A lost-sales production/inventory model with two discrete production modes', Stochastic Models, 16: 5, 453 — 478
To link to this Article: DOI: 10.1080/15326340008807600
URL: http://dx.doi.org/10.1080/15326340008807600

COMMUN. STATIST.—STOCHASTIC MODELS, 16(5), 453–478 (2000)

A LOST-SALES PRODUCTION/INVENTORY MODEL WITH TWO DISCRETE PRODUCTION MODES

Rein D. NOBEL* Mattijs van der HEEDEN

Department of Econometrics,
Faculty of Economics and Econometrics,
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands.

Abstract

A discrete production/inventory model is considered in which batch orders for a single item arrive at a production facility according to a Poisson process. The items can be produced according to two production modes, regular mode or high speed mode. Changing production mode requires a setup time during which production is disabled. Demand that cannot be satisfied from stock is lost. To control this model with respect to a suitable cost criterion, two-level hysteretic switching strategies are considered. Using generally applicable methods, tractable expressions are obtained for the fraction of lost demand and the average inventory level, amongst others. In these methods an essential role is played by the discrete Fast Fourier Transform.

*email: rnobel@econ.vu.nl

Keywords: production/inventory control, lost-sales, compound Poisson demand, embedded Markov chain, Fast Fourier Transform.

1 Introduction

Single-item production/inventory models in which the production rate can be adjusted to the on-hand inventory level in order to guarantee some prespecified service quality can play an important role in the inventory management of large industries. We study a discrete production model in which each item requires a positive production time, so only at so-called production completion epochs the inventory level increases. Orders arrive in batches according to a Poisson process and demand that cannot be satisfied upon arrival is lost. Of course there is a trade-off between, amongst others, holding costs for keeping items in inventory and the quality of service as expressed, e.g., in the fraction of lost demand. To enhance a fine tuning between the conflicting objectives of low holding costs and a small fraction of lost demand, the model allows for two production modes, regular mode and high speed mode. A change of mode can be initiated at production completion epochs only, and requires a switching time during which production is disabled. So, we are faced with a control problem to find the switching strategy that minimizes some over-all cost function, e.g. a weighted sum of the holding costs, operating costs and switching costs under the constraint that the fraction of lost demand does not exceed a prespecified acceptable percentage of lost-sales. In the analysis we will restrict ourselves to the natural class of so-called two-level hysteretic switching rules. Under such a control rule a change from regular mode to high speed mode takes place when the inventory level has dropped below a given threshold, and a change vice versa is initiated when at a production completion epoch the inventory has increased above some prespecified upper level. A precise description of the model will be given in Section 2.

Variants of this discrete production/inventory model with a two-critical-number policy have been widely studied for the case that the unsatisfied demand is backordered and the production is suspended when the inventory reaches the upper control level (see e.g. [2], [9], [12], [19], [20]). In practice these two model characteristics are not always the most adequate. It is quite natural to consider a regular production mode when the inventory level is high and a high speed mode when the inventory level is low. Also, the model assumption of backordering the unsatisfied demand often is more motivated by mathematical convenience than by practical considerations. Especially for so-called class B and C items (see [18]), considering unsatisfied demand as lost is more realistic. The lost-sales model is intrinsically more difficult (see below), and no exact results for the discrete model have been reported yet. Also for production/inventory models with a continuous production rate (see e.g. [6], [4]) attention has been paid mainly to the backorder case. The main objective of this paper is to present numerical methods which enable us to calculate all performance measures of interest for the lost-sales model with varying production modes and switch-over times.

The first step in our analysis is to obtain, for a fixed switching rule, the steady-state distribution of the Markov chain embedded at the production completion epochs. The discrete Fast Fourier Transform is essential to calculate the numerical values of the one-step transition probabilities of this Markov chain, and a geometric-tail property of the steady-state distribution is exploited to reduce the infinite set of equilibrium equations to a (small) finite set of linear equations. This Markov chain analysis is presented in Section 3.

The next, more difficult, step in the analysis of the lost-sales model is to find the average inventory level. The key idea for the solution of this problem is to express the conditional expectation of the cumulative inventory between two consecutive production completion epochs in terms of the coefficients of three explicit generating functions. The numerical values of these coefficients can routinely be obtained by the powerful discrete Fast Fourier Transform. Further details for finding the average inventory level are discussed in Section 4. This approach is new and seems generally applicable. In the Sections 5 and 6 a similar approach is used to obtain the long-run fraction of lost demand and the long-run fraction of customer orders that are not fully satisfied.

In Section 7 we briefly discuss the operating costs and the switching costs per unit time and, to conclude this paper, we describe the computation of the parameters of the two-level switching rule that solves the optimization problem as presented in Section 2. Numerical results are also presented. Comparison with extensive simulation experiments has shown that the procedures presented in this paper are numerically stable in a wide range of parameters.

We end this introduction with some general remarks to position the subject of this paper within the realm of queueing/inventory theory. The model presented here falls in the broad category of queueing and production/inventory models with state dependent parameters (see [7] for a recent overview). Queueing models with variable service and/or arrival rate that have been considered in the literature are often of the $M/G/1$-type with infinite waiting room (e.g. [8], [10], [13], [14], [15], [22] to cite only a few). Finite-buffer queueing systems with state dependent parameters are intrinsically more difficult to analyse and have received much less attention. A similar phenomenon can be noticed with respect to discrete production/inventory models with variable production rate. As mentioned above, the backorder case, which shows some parallelism with infinite-waiting-room queueing models, has been studied extensively, whereas tractable solutions for lost-sales models are often based on heuristics (e.g. see [5] for a continuous model with suspended production). The reason for the intractability of many finite-buffer systems and lost-sales models is rooted in the fact that their analysis leads to incomplete sums and integrals rather than to tractable complete sums and integrals, so paramount in many infinite-buffer models. And, not surprisingly, these incomplete sums and integrals cannot be reduced to simple expressions.

The merit of this paper is that we present a method to get round this problem by constructing several generating functions that have as their coefficients the different incomplete sums, present in the initial setup for the formulae of the performance measures. It turns out that these generating functions can often be evaluated explicitly, and so the incomplete sums themselves can be numerically calculated by inversion. Because nowadays the Fast Fourier Transform gives an excellent method for inverting generating functions (see e.g. [1]), we succeed to give numerical procedures for many steady-state performance measures of the lost-sales model presented in this paper. This method can also be used for many finite-buffer systems with state dependent parameters. See the forthcoming paper [16] for an example.

2 Description of the Model

We consider a single-item production/inventory model with two production modes. Orders for the item arrive at a production facility according to a Poisson process with rate λ. The number of items in an order, say B, has a general discrete probability distribution $\Pr\{B = k\} = \beta_k, k = 1, 2, \ldots$. The demand is satisfied directly upon arrival of an order from the inventory on hand. If the size of an order is larger than the inventory on hand, the excess demand is lost. The facility produces items without interruption according to one of two possible production modes, regular (R) or high speed (H) mode. The production time for an item in regular (high speed) mode, denoted by the generic random variable S_R (S_H), has a general distribution function $F_R(\cdot)$ ($F_H(\cdot)$). At production completion epochs only, the facility can decide to change production mode. If the facility decides to change mode, a setup time is involved to prepare for the new production mode. During a setup time production is disabled. The setup time to prepare the change from regular (high speed) to high speed (regular) mode is denoted by V_R (V_H) and has a general distribution function $G_R(\cdot)$ ($G_H(\cdot)$).

We impose four different types of costs on the model: holding, penalty, operating, and switching costs. More specifically, for each item held in stock a holding cost h per unit time is incurred. With each item lost (or, as an alternative, with every order not fully satisfied) we associate a penalty cost p. Thirdly, during regular (high speed) production periods an operating cost c_R (c_H) per unit time is incurred. And finally, each change in production mode involves a switching cost s. To balance these conflicting costs, we will now introduce an intuitively appealing control rule, that prescribes when to change production mode.

The rule we have in view is a so-called hysteretic (m, M) switching rule, where the control parameters m and M are integers with $0 \leq m < M$. Under this rule the facility switches from the regular (high speed) production mode to the other production mode when just after the completion of a production, done according to the regular (high speed) mode, the stock on hand is at or below level m (at level M). Here we count the newly produced item as not put on stock yet. In all other cases the next production starts immediately in the same mode as the last item produced.

This completes the description of the production/inventory model to be discussed in this paper. Of course, the main problem of interest with respect to this model is to find the switching rule that minimizes a suitable cost criterion. Although in practice penalty costs are difficult to ascertain and one usually formulates production control problems in a constrained setting, we will present our minimization problem in unconstrained form, using a Lagrangian approach, by including penalty costs for lost demand as introduced above.

To solve this problem, we will take a fixed (m, M)-switching rule as our starting point and concentrate on the calculation of the different performance measures of the resulting model (from now on, the (m, M)-system), which are present in the formulation of the control problem. Once we know these performance measures for any (m, M)-switching rule, the control problem can be solved by simple enumeration.

In the subsequent sections we will present numerical procedures for the following performance measures:

- $\eta_{m,M}$ = the long-run average inventory level,
- $\zeta_{m,M}$ = the long-run fraction of lost demand,
- $\kappa_{m,M}$ = the long-run fraction of orders not fully satisfied,
- $\rho_{m,M}^{(L)}$ = the long-run fraction of time that the production mode is in L-mode ($L = R, H$),
- $\sigma_{m,M}$ = the long-run average number of mode changes per unit time.

Our unconstrained optimization problem can now be formulated as follows,

$$\min_{m,M} \left\{ h\eta_{m,M} + c_R\rho^{(R)}_{m,M} + c_H\rho^{(H)}_{m,M} + s\sigma_{m,M} + p\zeta_{m,M}\lambda E[B] \right\}, \tag{2.1}$$

where h, c_R, c_H, s and p are given positive constants introduced above. As a side-remark, $\zeta_{m,M}\lambda E[B]$, the long-run average number of items lost per unit time, can be replaced by $\kappa_{m,M}\lambda$, if one considers penalties for not completely satisfied orders more realistic.

To guarantee steady-state behaviour, we require the stability condition, that under regular production mode return to inventory levels below m occurs with probability one, so the average demand per unit time must be larger than the regular production rate, i.e. $\lambda E[B]E[S_R] > 1$. Further, to make the model more realistic, we require of course $E[S_H] < E[S_R]$.

To conclude this section, we introduce some probabilities, that are the 'building blocks' for our calculations. Let

$$\begin{aligned} a_j^{(L)} &= \Pr\{ \text{ total demand during a production time } S_L \text{ is } j\} \\ \phi_j^{(L)} &= \Pr\{ \text{ total demand during a setup time } V_L \text{ is } j\} \end{aligned}$$

for $L = R, H$. The numerical values of these probabilities can be routinely computed from their generating functions by the discrete Fast Fourier Transform, see e.g. Section 1.2 in [21]. Denote by

$$\beta(z) = \sum_{n=1}^{\infty} \beta_n z^n$$

the generating function of the batch-size distribution $\{\beta_n, n \geq 1\}$. Then, it follows from a well-known result for the generating function of a compound Poisson variable that (see Section 1.3 in Tijms [21]) the required generating functions are given by,

$$A_L(z) := \sum_{j=0}^{\infty} a_j^{(L)} z^j = \hat{F}_L(\lambda(1-\beta(z))) \quad \text{for } L = H, R,$$

$$\Phi_L(z) := \sum_{j=0}^{\infty} \phi_j^{(L)} z^j = \hat{G}_L(\lambda(1-\beta(z))) \quad \text{for } L = H, R,$$

where $\hat{F}_L(\cdot)$ and $\hat{G}_L(\cdot)$ are the Laplace-Stieltjes transforms of $F_L(\cdot)$ and $G_L(\cdot)$ respectively. Also we will need the probabilities (again $L = R, H$),

$$\begin{aligned} \xi_j^{(L)} = \Pr\{ & \text{ total demand during the time from the start of a setup time } V_L \\ & \text{until the next production completion epoch is } j\}. \end{aligned}$$

Of course we have

$$\xi_j^{(L)} = \sum_{k=0}^{j} \phi_k^{(L)} a_{j-k}^{(L')}, \quad j \geq 0, \quad L \neq L'.$$

For many production-time and setup-time distributions of practical interest the above generating functions can be reduced to simple algebraic expressions. Ready-to-use codes for the discrete Fast Fourier Transform method are widely available, see e.g. Press et al. [17].

3 The Embedded Markov Chain Approach

Under the stability condition mentioned in Section 2, the continuous-time stochastic process $\{I_{m,M}(t), t \geq 0\}$, describing the inventory level in the (m, M)-system at time t, is a regenerative process for any (m, M) switching rule. We take as the regeneration points the high speed mode production completion epochs at which the stock on hand is at level M (as mentioned above, excluding the item just produced). From now on we consider the process $\{I_{m,M}(t)\}$ only for a fixed (m, M) switching strategy and suppress m and M in all future notations. We will study the process $\{I(t), t \geq 0\}$ by embedding a Markov chain at the production completion epochs. So, define

$$\begin{aligned} t_n &= \text{epoch of the } n\text{th production completion}, \\ L_n &= \text{production mode used for the production of the } n\text{th item}. \end{aligned}$$

For completeness, define $t_0 = 0$, $L_0 = H$ and $I(0+) = M + 1$, so the process starts in a regeneration point. Then, denoting by t_n- the epoch just before the nth production completion, the discrete-time process

$$X_n := (I(t_n-), L_n), \qquad n = 0, 1, \ldots,$$

is a time-homogeneous positive recurrent aperiodic Markov chain because the demand-process is Poisson and because of the above argument concerning the stability condition. The state space of this Markov chain is

$$\mathcal{S} = \{(j, R), j = 0, 1, \ldots\} \cup \{(j, H), j = 0, 1, \ldots, M\}.$$

Hence, the Markov chain $\{X_n\}$ has a limiting distribution which we denote by

$$\pi(j, L) := \lim_{n\to\infty} \Pr\{(I(t_n-), L_n) = (j, L)\} \qquad ((j, L) \in \mathcal{S}).$$

To calculate the limiting distribution $\{\pi(j, L)\}$ we first give the one-step transition probabilities,

$$p_{(i,K)(j,L)} = \Pr\{(I(t_{n+1}-), L_{n+1}) = (j, L) \mid (I(t_n-), L_n) = (i, K)\}.$$

For $j = 1, 2, \ldots$ we have with the notations of the previous section,

$$\begin{aligned} p_{(i,R)(j,R)} &= a^{(R)}_{i+1-j} \qquad i = m+1, \ldots;\ j = 1, \ldots, i+1, \\ p_{(i,R)(j,H)} &= \xi^{(R)}_{i+1-j} \qquad i = 0, \ldots, m;\ j = 1, \ldots, i+1, \\ p_{(i,H)(j,H)} &= a^{(H)}_{i+1-j} \qquad i = 0, \ldots, M-1;\ j = 1, \ldots, i+1, \\ p_{(M,H)(j,R)} &= \xi^{(H)}_{M+1-j} \qquad j = 1, \ldots, M+1. \end{aligned}$$

$$E[Y_n] = \sum_{(j,L)\in\mathcal{S}} y(j,L)p^{(n)}_{(M,H)(j,L)}, \tag{4.2}$$

$$E[Z_n] = \sum_{(j,L)\in\mathcal{S}} z(j,L)p^{(n)}_{(M,H)(j,L)}. \tag{4.3}$$

We can now formulate the following lemma.

Lemma 4.1

$$\lim_{n\to\infty} E[Y_n] = \sum_{(j,L)\in\mathcal{S}} y(j,L)\pi(j,L), \tag{4.4}$$

$$\lim_{n\to\infty} E[Z_n] = \sum_{(j,L)\in\mathcal{S}} z(j,L)\pi(j,L), \tag{4.5}$$

where $\{\pi(j,L)\}$ is again the limiting distribution of the Markov chain $\{X_n\}$.

Proof First we show that the right-hand sides (r.h.s.) of (4.4) and (4.5) are both finite. Since the $z(j,L)$ are bounded by the larger of $E[V_R]+E[S_H]$ and $E[V_H]+E[S_R]$ this is trivially true for (4.5). With respect to (4.4) we remark that it is sufficient to consider the infinite tail of the r.h.s., i.e.

$$\sum_{j=M+1}^{\infty} y(j,R)\pi(j,R). \tag{4.6}$$

Since it can be easily seen that for all $j = M+1, M+2, \ldots$

$$y(j,R) \leq (j+1)E[S_R]$$

and we know that

$$\pi(j,R) = \frac{\pi(M+1)}{\tau^{M+1}}\tau^j,$$

we can conclude that also (4.6) is finite.

Next, as in [11], we note that from the n-step balance equations

$$\pi(j,L) = \sum_{(i,L')\in\mathcal{S}} \pi(i,L')p^{(n)}_{(i,L')(j;L)}$$

it follows, by taking only one term in the r.h.s., e.g. $(i,L') = (M,H)$, that

$$p^{(n)}_{(M,H)(j,L)} \leq \frac{\pi(j,L)}{\pi(M,H)}.$$

So we can conclude that the terms of the r.h.s. of (4.2) and (4.3) are bounded by

$$\frac{y(j,L)}{\pi(M,H)}\pi(j,L) \quad \text{and} \quad \frac{z(j,L)}{\pi(M,H)}\pi(j,L),$$

respectively. The lemma follows by applying the bounded convergence theorem. □

We will now consider the general state space Markov chain $\{W_n\}$. In view of our stability condition (recall, $\lambda E[B]E[S_R] > 1$) the Markov chain $\{W_n\}$ is an aperiodic positive recurrent Harris chain (shortly, Harris ergodic), because it has as a regeneration set

$$\Sigma_0 = \{(x, y, z) \in \Sigma : x = (M, H), y \geq 0, z \geq 0\},$$

and the number of transitions between two visits to Σ_0 is aperiodic and has finite expectation (see e.g. [3]). As a consequence, the Markov chain $\{W_n\}$ has a stationary distribution, say ν. Now, introduce for convenience the random vector (X, Y, Z) with its simultaneous distribution given by ν. Then, the marginal distribution of X is given by the distribution $\{\pi(j, L)\}$ and we can state the following result.

Theorem 4.1 *The long-run average inventory η is given by*

$$\eta = \frac{E[Y]}{E[Z]} = \frac{\sum_{(j,L)\in\mathcal{S}} y(j, L)\pi(j, L)}{\sum_{(j,L)\in\mathcal{S}} z(j, L)\pi(j, L)}$$

with probability one.

Proof From (4.1) we see that it is sufficient to justify the application of the ergodic theorem, for both the numerator and the denominator, i.e. for $N \to \infty$,

$$\frac{1}{N}\sum_{n=0}^{N-1} Y_n \to E[Y],$$

with probability one (and similarly for $E[Z]$). We only discuss the numerator. Since we know from the proof of Lemma 4.1 that

$$E[Y] = \sum_{(j,L)\in\mathcal{S}} y(j, L)\pi(j, L)$$

is finite, we can use the Harris ergodicity of $\{W_n\}$ to conclude that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} Y_n = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} E[Y_n], \quad \text{w. p. 1.}$$

And so the proof is complete, because, again using Lemma 4.1, the Cesaro-limit on the right is equal to $E[Y]$. □

So, the calculation of η boils down to the question how to determine $E[Y]$ and $E[Z]$ via the $y(j, L)$ and the $z(j, L)$, respectively. Concerning $E[Z]$ the situation is rather simple, because the $z(j, L)$ are simple expectations. Using also the geometric-tail behaviour of the $\{\pi(j, R)\}$ for $j \geq M + 1$, we get

$$E[Z] = \sum_{(j,L)\in\mathcal{S}} z(j, L)\pi(j, L) = E[S_H]\sum_{j=0}^{M-1} \pi(j, H) + \{E[V_H] + E[S_R]\}\pi(M, H)$$

$$+\{E[V_R]+E[S_H]\}\sum_{j=0}^{m}\pi(j,R)+E[S_R]\left\{\sum_{j=m+1}^{M}\pi(j,R)+\frac{\pi(M+1,R)}{1-\tau}\right\}. \tag{4.7}$$

To calculate $E[Y]$ we need the $y(j,L)$ and, as for the $z(j,L)$, there are four different cases,

$$\begin{aligned}
y(j,R) &= E\left[\int_0^{V_R+S_H} I(t)\,dt \middle| I(0)=j+1\right], \quad j=0,\dots,m \\
y(j,R) &= E\left[\int_0^{S_R} I(t)\,dt \middle| I(0)=j+1\right], \quad j=m+1,\dots, \\
y(j,H) &= E\left[\int_0^{S_H} I(t)\,dt \middle| I(0)=j+1\right], \quad j=0,\dots,M-1. \\
y(M,H) &= E\left[\int_0^{V_H+S_R} I(t)\,dt \middle| I(0)=M+1\right].
\end{aligned}$$

So, in words, we need an algorithm to calculate the expected area under the graph of the inventory function $I(\cdot)$ between two consecutive production completion epochs, given the inventory level at the first epoch.

The following theorem presents a key result for the calculation of these conditional expectations.

Theorem 4.2 *Let the generic non-negative random variable V denote a time lapse between two consecutive production completion epochs with distribution function $G(\cdot)$ and Laplace-Stieltjes transform $\hat{G}(\cdot)$. Consider the following generating functions,*

$$\begin{aligned}
\Gamma_r(z) &= \sum_{k=0}^{\infty}\gamma_k(r)z^k = -r\hat{G}'(\lambda(1-\beta(z)))-\frac{1}{2}\lambda z\beta'(z)\hat{G}''(\lambda(1-\beta(z))), \qquad r\geq 1, \\
\Delta(z) &= \sum_{k=0}^{\infty}\delta_k z^k = \frac{z}{(1-z)(1-\beta(z))}\left\{\frac{1}{\lambda}\left(1-\hat{G}(\lambda(1-\beta(z)))\right)+(1-\beta(z))\hat{G}'(\lambda(1-\beta(z)))\right\}, \\
\Theta(z) &= \sum_{k=0}^{\infty}\theta_k z^k = \frac{z\beta'(z)}{(1-z)(1-\beta(z))^2} \\
&\quad\times\left\{\frac{1}{\lambda}\left(1-\hat{G}(\lambda(1-\beta(z)))\right)+(1-\beta(z))\hat{G}'(\lambda(1-\beta(z)))-\frac{1}{2}\lambda(1-\beta(z))^2\hat{G}''(\lambda(1-\beta(z)))\right\}.
\end{aligned}$$

Then, for any $r\geq 1$,

$$E\left[\int_0^V I(t)\,dt\middle| I(0)=r\right]=\sum_{k=0}^{r-1}\gamma_k(r)+r\delta_r-\theta_{r-1}.$$

For the proof, see Lemma 4.2–4.4 and the Appendix below.

With this result we are able to find expressions for the conditional expectations $y(j,L)$ in all the four cases presented above by taking the random variable V equal to V_R+S_H, S_R, S_H and V_H+S_R respectively, or, in other words, by choosing the Laplace-Stieltjes transform $\hat{G}(\cdot)$ equal to $\hat{G}_R(\cdot)\hat{F}_H(\cdot)$, $\hat{F}_R(\cdot)$, $\hat{F}_R(\cdot)$ and $\hat{G}_R(\cdot)\hat{F}_H(\cdot)$. Adding self-explanatory superscripts to all the coefficients to distinguish the different cases and exploiting again the geometric-tail behaviour of the $\pi(j,R)$ for $j\geq M+1$, we can calculate $E[Y]$ as follows,

$$
\begin{aligned}
E[Y] &= \sum_{(j,L)\in\mathcal{S}} y(j,L)\pi(j,L) = \\
&\sum_{j=0}^{M-1}\left[\sum_{k=0}^{j}\gamma_k^{(H)}(j+1)+(j+1)\delta_{j+1}^{(H)}-\theta_j^{(H)}\right]\pi(j,H) \\
&+\left[\sum_{k=0}^{M}\gamma_k^{(HR)}(M+1)+(M+1)\delta_{M+1}^{(HR)}-\theta_M^{(HR)}\right]\pi(M,H) \\
&+\sum_{j=0}^{m}\left[\sum_{k=0}^{j}\gamma_k^{(RH)}(j+1)+(j+1)\delta_{j+1}^{(RH)}-\theta_j^{(RH)}\right]\pi(j,R) \qquad (4.8)\\
&+\sum_{j=m+1}^{M}\left[\sum_{k=0}^{j}\gamma_k^{(R)}(j+1)+(j+1)\delta_{j+1}^{(R)}-\theta_j^{(R)}\right]\pi(j,R) \\
&+\frac{\pi(M+1,R)}{\tau^{M+1}}\sum_{j=M+1}^{\infty}\left[\sum_{k=0}^{j}\gamma_k^{(R)}(j+1)+(j+1)\delta_{j+1}^{(R)}-\theta_j^{(R)}\right]\tau^j.
\end{aligned}
$$

Notice that, as for $E[Z]$, we can get rid of the infinite sum in the last term. We first rewrite the complete infinite series, i.e. from $j=0$, as follows

$$
\sum_{j=0}^{\infty}\left[\sum_{k=0}^{j}\gamma_k^{(R)}(j+1)+(j+1)\delta_{j+1}^{(R)}-\theta_j^{(R)}\right]\tau^j = \frac{\lambda\tau\beta'(\tau)A_R''(\tau)}{2(1-\tau)}-\frac{A_R'(\tau)}{(1-\tau)^2}+\frac{d}{dz}\left[\Delta^{(R)}(z)\right]_{z=\tau}-\Theta^{(R)}(\tau), \qquad (4.9)
$$

and subtract subsequently the finite sum

$$
\sum_{j=0}^{M}\left[\sum_{k=0}^{j}\gamma_k^{(R)}(j+1)+(j+1)\delta_{j+1}^{(R)}-\theta_j^{(R)}\right]\tau^j.
$$

To find the closed form expression for the infinite series (4.9) requires only some simple algebra, so we skip any further details.

In conclusion, to find $E[Y]$ we only need to apply Theorem 4.2 for different values of r and random variables V. As a final remark, the coefficients $\gamma_k(r)$, δ_k and θ_k are (not surprisingly) calculated by various applications of the discrete Fast Fourier Transform method.

The rest of this section will be dedicated to the proof of Theorem 4.2. To enhance readability we split the argument in a series of lemmas, but we first introduce the necessary notations.

$$
\begin{aligned}
X_j &= \text{time lapse between the epochs of the } (j-1)\text{th and the } j\text{th order arrival,}\\
N(t) &= \text{the number of orders arrived up to time } t,\\
B(t) &= \text{the total demand arrived up to time } t,\\
B_i &= \text{the size of the } i\text{th order.}
\end{aligned}
$$

Further, we need the n-fold convolution of the batch size distribution $\{\beta_k\}$,

$$
\beta_k^{*n} = \Pr\left\{\sum_{i=1}^{n} B_i = k\right\},
$$

with the usual conventions, $\beta_0^{*0}=1$ and $\beta_k^{*0}=0$ for $k\neq 0$, and the conditional probabilities

$$\beta({}^{*j}_{b}|{}^{*n}_{k}) := \Pr\left\{\sum_{i=1}^{j} B_i = b \,\middle|\, \sum_{i=1}^{n} B_i = k\right\}, \quad j \le n,$$

with the convention $\beta({}^{*0}_{0}|{}^{*0}_{0}) = 1$. Remark that $\beta({}^{*j}_{b}|{}^{*n}_{k}) = 0$ for $b < j$, because $\beta_0 = 0$. We now formulate three lemmas which jointly prove Theorem 4.2.

Lemma 4.2 *With the notation of Theorem 4.2,*

$$E\left[\int_0^V I(t)\,dt \,\middle|\, I(0) = r\right] =$$
$$\sum_{k=0}^{r-1}\sum_{n=0}^{k}\int_0^\infty \sum_{j=0}^{n}\frac{t}{n+1}\sum_{b=j}^{k}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)$$
$$+\sum_{k=r}^{\infty}\sum_{n=0}^{k}\int_0^\infty \sum_{j=0}^{n}\frac{t}{n+1}\sum_{b=j}^{r-1}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Lemma 4.3 *For the $\gamma_k(r)$, introduced in Theorem 4.2, we have for $k \ge 0$ and $r \ge 1$,*

$$\gamma_k(r) = \sum_{n=0}^{k}\int_0^\infty \sum_{j=0}^{n}\frac{t}{n+1}\sum_{b=j}^{k}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Lemma 4.4 *With the notation introduced in Theorem 4.2 we have for $r \ge 1$,*

$$\sum_{k=r}^{\infty}\sum_{n=0}^{k}\int_0^\infty \sum_{j=0}^{n}\frac{t}{n+1}\sum_{b=j}^{r-1}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t) = r\delta_r - \theta_{r-1}.$$

We see that Theorem 4.2 follows directly from these three lemmas. In this section we give only the proof of Lemma 4.2. The other, rather technical, proofs are given in the Appendix.

Proof of Lemma 4.2: We will use conditioning on the random variable V, the number $N(V)$ of order arrivals during V and the total demand $B(V)$ during V. Then we get

$$E\left[\int_0^V I(t)\,dt \,\middle|\, I(0) = r\right] =$$
$$\int_0^\infty \sum_{k=0}^{\infty}\sum_{n=0}^{k}\sum_{j=0}^{n} E\left[\left(r - \sum_{i=1}^{j} B_i\right)^+ X_{j+1} \,\middle|\, V = t; B(V) = k; N(V) = n\right]$$
$$\times \Pr\{B(V) = k; N(V) = n | V = t\}\,dG(t), \tag{4.10}$$

where for convenience $X_{n+1} := t - \sum_{j=1}^{n} X_j$ and the other X_j as defined above. Now, by the independence between V, the arrival process of the orders $\{N(t)\}$ and the order sizes $\{B_i\}$, and using the well-known property of the Poisson process that

$$E[X_j | N(t) = n] = \frac{t}{n+1}, \quad j = 1, \ldots, n+1,$$

expression (4.10) can be rewritten as,

$$\int_0^\infty \sum_{k=0}^{\infty}\sum_{n=0}^{k}\sum_{j=0}^{n} E\left[\left(r-\sum_{i=1}^{j}B_i\right)^+ X_{j+1}\middle| B(t)=k; N(t)=n\right]\Pr\{B(t)=k;N(t)=n\}\,dG(t) =$$

$$\int_0^\infty \sum_{k=0}^{\infty}\sum_{n=0}^{k}\sum_{j=0}^{n} \frac{t}{n+1} E\left[\left(r-\sum_{i=1}^{j}B_i\right)^+ \middle| \sum_{i=1}^{n}B_i=k; N(t)=n\right]\Pr\left\{\sum_{i=1}^{n}B_i=k;N(t)=n\right\}\,dG(t) =$$

$$\int_0^\infty \sum_{k=0}^{\infty}\sum_{n=0}^{k}\sum_{j=0}^{n} \frac{t}{n+1} E\left[\left(r-\sum_{i=1}^{j}B_i\right)^+ \middle| \sum_{i=1}^{n}B_i=k\right]\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t) =$$

$$\int_0^\infty \sum_{k=0}^{r-1}\sum_{n=0}^{k}\sum_{j=0}^{n} \frac{t}{n+1}\sum_{b=j}^{k}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)$$

$$+\int_0^\infty \sum_{k=r}^{\infty}\sum_{n=0}^{k}\sum_{j=0}^{n} \frac{t}{n+1}\sum_{b=j}^{r-1}(r-b)\beta({}^{*j}_{b}|{}^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

□

5 The Fraction of Lost Demand

Our next performance measure of interest is the long-run fraction of lost demand,

$$\zeta_{m,M} := \lim_{t\to\infty}\frac{L_{m,M}(t)}{B(t)},$$

where $L_{m,M}(t)$ is the total demand lost up to time t. Again we fix a (m, M) switching rule and suppress the subscripts. Define

$Q_n = L(t_{n+1}) - L(t_n) =$

the number of items lost between the nth and the $(n+1)$th production completion epoch,

$D_n = B(t_{n+1}) - B(t_n) =$

the total demand between the nth and the $(n+1)$th production completion epoch.

Then ζ can be rewritten as

$$\zeta = \lim_{N\to\infty}\frac{\sum_{n=0}^{N-1}Q_n}{\sum_{n=0}^{N-1}D_n} = \lim_{N\to\infty}\frac{\frac{1}{N}\sum_{n=0}^{N-1}Q_n}{\frac{1}{N}\sum_{n=0}^{N-1}D_n}. \tag{5.1}$$

Again we will show that both the numerator and the denominator of (5.1) converge to a constant value with probability one. To do this, we remark that the stochastic process $U_n := (X_{n+1}, D_n, Q_n)$, is an aperiodic positive recurrent Markov chain with countable state space $\mathcal{S} \times \{0,1,\ldots\} \times \{0,1,\ldots\}$. This Markov chain has the property that the distribution of U_n depends on U_{n-1} only through the first coordinate, i.e. X_n. So, we can proceed as in Section 4. Let the random vector (X, D, Q) have the limiting distribution of the Markov chain $\{U_n\}$. Then, starting the Markov chain $\{X_n\}$ in $X_0 = (M, H)$, and defining

$$q(j,L)=E[Q_n \mid X_n=(j,L)] \quad \text{and} \quad d(j,L)=E[D_n \mid X_n=(j,L)]$$

we have, using the n-step transition probabilities $p^{(n)}_{(M,H)(j,L)}$,

$$\begin{aligned} E[Q_n] &= \sum_{(j,L)\in\mathcal{S}} q(j,L)p^{(n)}_{(M,H)(j,L)}, \\ E[D_n] &= \sum_{(j,L)\in\mathcal{S}} d(j,L)p^{(n)}_{(M,H)(j,L)}. \end{aligned}$$

Now, exactly as in Section 4, we can show that the $E[Q_n]$ and the $E[D_n]$ converge to $E[Q]$ and $E[D]$ respectively, and the ergodic theorem gives

$$\zeta=\lim_{N\to\infty}\frac{\frac{1}{N}\sum_{n=0}^{N-1}Q_n}{\frac{1}{N}\sum_{n=0}^{N-1}D_n}=\frac{E[Q]}{E[D]} \quad \text{w.p.1.}$$

Hence, the calculation of ζ is reduced to the calculation of $E[Q]$ and $E[D]$.

The calculation of $E[Q]$ is straightforward, because

$$E[Q]=\sum_{(j,L)\in\mathcal{S}} q(j,L)\pi(j,L) \tag{5.2}$$

with the $q(j,L)$ given as follows

$$\begin{aligned} q(j,R) &= \sum_{k>j+1}(k-j-1)\xi^{(R)}_k, \quad j=0,\dots,m \\ q(j,R) &= \sum_{k>j+1}(k-j-1)a^{(R)}_k, \quad j=m+1,\dots, \\ q(j,H) &= \sum_{k>j+1}(k-j-1)a^{(H)}_k, \quad j=0,\dots,M-1. \\ q(M,H) &= \sum_{k>M+1}(k-M-1)\xi^{(H)}_k. \end{aligned}$$

With some simple algebra all these quantities $q(j,L)$ can be expressed as finite sums, and, using the geometric-tail behaviour of the $\pi(j,R)$, also the infinite tail $\sum_{j>M} q(j,R)\pi(j,R)$ in (5.2) can be rewritten as an expression in which only finite sums occur. So we have an efficient algorithm for $E[Q]$. For completeness, we give the final result

$$\begin{aligned} E[Q] = &\sum_{(j,L)\in\mathcal{S}} q(j,L)\pi(j,L) = \\ &\sum_{j=0}^{M-1}\left[\lambda E[B]E[S_H]-\sum_{k=1}^{j+1}ka^{(H)}_k-(j+1)\left(1-\sum_{k=0}^{j+1}a^{(H)}_k\right)\right]\pi(j,H) \\ &+\left[\lambda E[B](E[V_H]+E[S_R])-\sum_{k=1}^{M+1}k\xi^{(H)}_k-(M+1)\left(1-\sum_{k=0}^{M+1}\xi^{(H)}_k\right)\right]\pi(M,H) \\ &+\sum_{j=0}^{m}\left[\lambda E[B](E[V_R]+E[S_H])-\sum_{k=1}^{j+1}k\xi^{(R)}_k-(j+1)\left(1-\sum_{k=0}^{j+1}\xi^{(R)}_k\right)\right]\pi(j,R) \\ &+\sum_{j=m+1}^{M}\left[\lambda E[B]E[S_R]-\sum_{k=1}^{j+1}ka^{(R)}_k-(j+1)\left(1-\sum_{k=0}^{j+1}a^{(R)}_k\right)\right]\pi(j,R) \\ &+\frac{\pi(M+1,R)}{\tau^{M+1}}\left\{\frac{\tau^{M+1}\lambda E[B]E[S_R]-1}{1-\tau}+\sum_{j=0}^{M}\left[\sum_{k=1}^{j+1}ka^{(R)}_k+(j+1)\left(1-\sum_{k=0}^{j+1}a^{(R)}_k\right)\right]\tau^j\right\}. \end{aligned} \tag{5.3}$$

The calculation of $E[D]$ is even simpler because

$$E[D] = \sum_{(j,L)\in\mathcal{S}} d(j,L)\pi(j,L)$$

and the compound Poisson demand gives $d(j,L) = \lambda E[B]z(j,L)$, with $z(j,L)$ as introduced in Section 4. So we can conclude that

$$E[D] = \lambda E[B]E[Z],$$

where for $E[Z]$ we use expression (4.7).

6 The Fraction of Orders Not Completely Satisfied

The next performance measure we consider is the long-run fraction of orders not completely satisfied,

$$\kappa_{m,M} = \lim_{t\to\infty} \frac{J_{m,M}(t)}{N(t)},$$

where $J_{m,M}(t)$ is the number of orders arriving in $(0,t)$ that are not completely satisfied. Fixing a strategy (m,M), suppressing indices, and defining

$$K_n := J(t_{n+1}) - J(t_n),$$

we can rewrite κ as

$$\kappa = \lim_{N\to\infty} \frac{\frac{1}{N}\sum_{n=0}^{N-1} K_n}{\frac{1}{N}\sum_{n=0}^{N-1}(N(t_{n+1}) - N(t_n))}.$$

Argueing as in the previous sections, we can see that the denominator converges to $\lambda E[Z]$ and the numerator to a constant which again can be considered as an expectation $E[K]$, where K has the limiting distribution

$$\Pr(K = i) = \lim_{n\to\infty} \Pr(K_n = i),$$

which is the marginal distribution of the limiting distribution of the Markov chain $\{(X_{n+1}, K_n)\}$. Introducing

$$k(j,L) = E[K_n \mid X_n = (j,L)]$$

the expectation $E[K]$ will be calculated by

$$E[K] = \sum_{(j,L)\in\mathcal{S}} k(j,L)\pi(j,L),$$

where the $k(j,L)$ can take the following forms

$$k(j,R) \quad = \quad E[J(V_R + S_H) \mid I(0) = j+1], \quad j = 0,\ldots,m$$

$$\begin{aligned} k(j,R) &= E[J(S_R) \mid I(0)=j+1], \quad j=m+1,\ldots, \\ k(j,H) &= E[J(S_H) \mid I(0)=j+1], \quad j=0,\ldots,M-1, \\ k(M,H) &= E[J(V_H+S_R) \mid I(0)=M+1]. \end{aligned}$$

To calculate the $k(j,L)$ we need an analogue for Theorem 4.2.

Theorem 6.1 *Let the generic non-negative random variable V denote a time lapse between two consecutive production completion epochs with distribution function $G(\cdot)$ and Laplace-Stieltjes transform $\hat{G}(\cdot)$. Consider the following generating functions,*

$$\Psi(z) := \sum_{k=0}^{\infty} \psi_k z^k = \hat{G}(\lambda(1-\beta(z)))$$

and

$$\Upsilon(z) := \sum_{k=0}^{\infty} \upsilon_k z^k = \frac{1}{1-z}\left\{E[V] + \frac{\hat{G}(\lambda(1-\beta(z)))-1}{\lambda(1-\beta(z))}\right\}.$$

Then, for any $r \geq 0$,

$$E[J(V) \mid I(0)=r] = \sum_{k=r+1}^{\infty} \psi_k + \lambda \upsilon_r.$$

The proof will be given below.

Using Theorem 6.1, we can calculate all the different conditional expectations $k(j,L)$ mentioned above, because, in complete analogy with Section 4, we only need to apply Theorem 6.1 for different values of r and random variables V. For completeness we give the expression for $E[K]$, where we remark that again the infinite tail $\sum_{j>M} k(j,R)\pi(j,R)$ has been rewritten using the geometric-tail behaviour of the $\pi(j,R)$, but we will not give the details (the superscripts in the υ-coefficients have of course the same meaning as in Section 4)

$$\begin{aligned} E[K] = &\sum_{(j,L)\in S} k(j,L)\pi(j,L) = \\ &\sum_{j=0}^{M-1}\left[\sum_{k=j+2}^{\infty} a_k^{(H)} + \lambda \upsilon_{j+1}^{(H)}\right]\pi(j,H) + \left[\sum_{k=M+2}^{\infty} \xi_k^{(H)} + \lambda \upsilon_{M+1}^{(HR)}\right]\pi(M,H) \\ &+\sum_{j=0}^{m}\left[\sum_{k=j+2}^{\infty} \xi_k^{(R)} + \lambda \upsilon_{j+1}^{(RH)}\right]\pi(j,R) + \sum_{j=m+1}^{M}\left[\sum_{k=j+2}^{\infty} a_k^{(R)} + \lambda \upsilon_{j+1}^{(R)}\right]\pi(j,R) \\ &+\frac{\pi(M+1,R)}{\tau^{M+1}}\left\{\frac{\lambda E[S_R]}{1-\tau} - \frac{\beta(\tau)}{\tau(1-\beta(\tau))} - \sum_{j=0}^{M}\left[\sum_{k=j+2}^{\infty} a_k^{(R)} + \lambda \upsilon_{j+1}^{(R)}\right]\tau^j\right\}. \end{aligned} \tag{6.1}$$

Proof of Theorem 6.1: Define

Y_r = time needed to run out of stock, given that at epoch 0 the inventory level is r

('run out of stock' means that at least one item has got lost). Then it can be easily seen that

$$E[J(V) \mid I(0)=r] = \Pr\{Y_r < V\} + \lambda E[(V-Y_r)^+].$$

Because (cf. Section 2)

$$\psi_k = \Pr\{\text{total demand during time } V \text{ is } k\},$$

we have

$$\Pr\{Y_r < V\} = \sum_{k=r+1}^{\infty} \psi_k.$$

So, to complete the proof of the theorem, we still have to show that

$$v_r = E[(V - Y_r)^+].$$

To do this we need the probability distribution function of the random variable Y_r. Let $y \geq 0$ and observe that $\{Y_r \leq y\} = \{B(y) > r\}$. This gives, by conditioning on $N(y)$, and independence arguments,

$$\begin{aligned}\Pr\{Y_r \leq y\} &= \sum_{k=r+1}^{\infty} \Pr\{B(y) = k\} = \sum_{k=r+1}^{\infty} \sum_{n=1}^{k} \Pr\{B(y) = k; N(y) = n\} = \\ &\sum_{k=r+1}^{\infty} \sum_{n=1}^{k} \Pr\left\{\sum_{i=1}^{n} B_i = k; N(y) = n\right\} = \sum_{k=r+1}^{\infty} \sum_{n=1}^{k} \beta_k^{*n} e^{-\lambda y} \frac{(\lambda y)^n}{n!}. \qquad (6.2)\end{aligned}$$

Next, using (6.2),

$$\begin{aligned}\sum_{r=0}^{\infty} E[(V - Y_r)^+] z^r &= \sum_{r=0}^{\infty} \int_0^{\infty} \int_0^t (t - y)\, d\Pr\{Y_r \leq y\}\, dG(t) z^r = \\ &\int_0^{\infty} \int_0^t (t - y)\, d\left(\sum_{r=0}^{\infty} \sum_{k=r+1}^{\infty} \sum_{n=1}^{k} \beta_k^{*n} e^{-\lambda y} \frac{(\lambda y)^n}{n!} z^r\right) dG(t).\end{aligned}$$

By interchanging the summations, we can evaluate the series between the big parentheses to

$$\frac{1}{1-z}\left(1 - e^{-\lambda y(1-\beta(z))}\right),$$

so we can also evaluate the double integral, and this gives the result. □

7 Numerical Results

Returning our attention to the optimization problem as formulated in (2.1), we see that still three quantities have to be calculated, i.e. $\rho_{m,M}^{(R)}$, $\rho_{m,M}^{(H)}$ and $\sigma_{m,M}$. Using the same type of arguments as in the previous sections, it is now straightforward to find expressions for these quantities. We give only the final results, without any further comment,

$$\begin{aligned}\rho_{m,M}^{(R)} &= \frac{E[S_R]}{E[Z]}\left(\pi(M, H) + \sum_{j=m+1}^{M} \pi(j, R) + \frac{\pi(M+1, R)}{1-\tau}\right), \\ \rho_{m,M}^{(H)} &= \frac{E[S_H]}{E[Z]}\left(\sum_{j=0}^{M-1} \pi(j, H) + \sum_{j=0}^{m} \pi(j, R)\right),\end{aligned}$$

$$\sigma_{m,M} = \frac{1}{E[Z]}\left(\pi(M,H)+\sum_{j=0}^{m}\pi(j,R)\right).$$

Now, reintroducing subscripts for all quantities dependent on the (m, M) switching rule, we can write down the definitive expression for the criterion function,

$$\begin{aligned} &h\eta_{m,M} + c_R\rho^{(R)}_{m,M} + c_H\rho^{(H)}_{m,M} + s\sigma_{m,M} + \zeta_{m,M}\lambda E[B] = \\ &\quad \frac{1}{E[Z_{m,M}]}\left\{hE[Y_{m,M}] + c_R E[S_R]\left(\pi(M,H) + \sum_{j=m+1}^{M}\pi(j,R) + \frac{\pi(M+1,R)}{1-\tau}\right)\right. \qquad (7.1)\\ &\quad \left. + c_H E[S_H]\left(\sum_{j=0}^{M-1}\pi(j,H) + \sum_{j=0}^{m}\pi(j,R)\right) + s\left(\pi(M,H) + \sum_{j=0}^{m}\pi(j,R)\right) + pE[Q_{m,M}]\right\}. \end{aligned}$$

Next, we can solve (2.1) by simple enumeration: start with some large $M = M_0$ and calculate for $m = 0, 1, \ldots, M-1$ the right hand side of (7.1). In this way we find the optimal switching strategy with $M = M_0$. Repeat this procedure for $M = M_0 - 1, M_0 - 2, \ldots,$ until $M = 1$. Of course, for every (m, M) strategy we need to solve the (small) system of $2M + 3$ balance equations (3.2), (3.5) and (3.6) to find the distribution $\{\pi(j, L)\}$ and we have to calculate the quantities $E[Z_{m,M}]$, $E[Y_{m,M}]$ and $E[Q_{m,M}]$, using the expressions (4.7), (4.8), and (5.3), respectively, but we emphasize that the number τ and all quantities that require an application of the discrete Fast Fourier Tranform method have to be evaluated only once. So, after all, the enumeration algorithm is reasonably efficient.

To illustrate the numerical procedures discussed in this paper, we will present the optimal strategies w.r.t. the chosen criterion function, together with the corresponding performance measures, for varying switching times and varying high speed production times. In all numerical examples we have kept the following parameters constant,

$$\lambda = 0.3, \quad E[B] = 4, \quad E[S_R] = 0.9.$$

We have taken a Coxian-2 distribution for the regular production time with squared coefficient of variation $c^2_{S_R} = 0.8$. Further, the cost parameters are always taken as follows,

$$h = 0.01, \quad p = 7.5, \quad c_R = 0.7, \quad c_H = 0.8, \quad s = 5.$$

First, the switching times V_R and V_H are varied, but always taken deterministic and equal. In Table 1 we show the optimal strategies for a geometric batch size (note that for this case $\zeta_{m,M} = \kappa_{m,M}$) and in Table 2 the corresponding values for a constant batch size. In all cases the high speed production time S_H has been given a Coxian-2 distribution with $E[S_H] = 0.8$ and squared coefficient of variation $c^2_{S_H} = 2$. Next, for switching times V_R and V_H deterministic and both equal to 1, we varied the expected high speed production time $E[S_H]$, keeping $c^2_{S_H} = 2$

Table 1: Optimal strategies for geometric batch size and varying constant switching times

$E[V_R]$	(m, M)	$\eta_{m,M}$	$\sigma_{m,M}$	$\rho^R_{m,M}$	$\rho^H_{m,M}$	$\zeta_{m,M} = \kappa_{m,M}$	Criterion
0	(18,49)	63.1069	0.003416	0.67185	0.32815	0.036093	1.7058
1	(21,53)	63.6840	0.003277	0.64082	0.35591	0.035915	1.7098
2	(24,56)	63.9199	0.003243	0.61234	0.38118	0.035963	1.7127
3	(29,59)	64.3281	0.003373	0.57622	0.41366	0.035567	1.7145
4	(50,61)	61.4182	0.007149	0.41907	0.55233	0.036624	1.7148
5	(77,78)	51.116	0.012790	0.10769	0.82836	0.037411	1.6499

Table 2: Optimal strategies for constant batch size and varying constant switching times

$E[V_R]$	(m, M)	$\eta_{m,M}$	$\sigma_{m,M}$	$\rho^R_{m,M}$	$\rho^H_{m,M}$	$\zeta_{m,M}$	$\kappa_{m,M}$	Criterion
0	(17,45)	45.3906	0.003629	0.61153	0.38847	0.029112	0.044001	1.4729
1	(19,48)	45.7194	0.003471	0.58098	0.41555	0.029193	0.044122	1.4764
2	(22,51)	46.2537	0.003411	0.54790	0.44528	0.028855	0.043612	1.4790
3	(25,53)	46.3476	0.003478	0.51563	0.47393	0.028883	0.043653	1.4809
4	(30,55)	46.5141	0.003777	0.47073	0.51416	0.028555	0.043159	1.4819
5	(64,65)	38.0818	0.01249	0.05307	0.88449	0.029521	0.044618	1.4537

in all cases. The results for a geometric batch size are given in Table 3, and the corresponding results for a constant batch size in Table 4. From these numerical results we see that the optimal strategies are rather insensitive for the batch size distribution, whereas the performance measures are quite different. Further, we can conclude that increasing the expected switching times or the high speed production times leads to higher optimal switching levels m and M. The average inventory, the number of mode changes per unit time, and the fraction of lost demand turn out to be not very sensitive for the expected switching times. Only when the switching times become very large we see an increase of the number of mode changes, simultaneously with a lower inventory level and a shift towards a higher fraction of time that the high speed production mode is used.

We see from Table 3 and 4 that, as the expected high speed production times $E[S_H]$ increase, the average inventory, the fraction of lost demand, the number of mode changes, and the fraction of time that the high speed production mode is used all increase. Only when $E[S_H]$ approaches $E[S_R]$ the average inventory level decreases at the cost of a higher fraction of lost demand.

8 Appendix.

In this appendix we prove the technical Lemmas 4.3 and 4.4 of Section 4.

Table 3: Optimal strategies for geometric batch size and varying $E[S_H]$

$E[S_H]$	(m,M)	$\eta_{m,M}$	$\sigma_{m,M}$	$\rho^R_{m,M}$	$\rho^H_{m,M}$	$\zeta_{m,M}=\kappa_{m,M}$	Criterion
0.1	(7,22)	63.7064	0.009803	0.97981	0.01039	0.006222	1.4363
0.2	(8,23)	64.0053	0.009672	0.96672	0.02361	0.006531	1.4428
0.3	(9,25)	64.6225	0.009002	0.95034	0.04066	0.007116	1.4531
0.4	(10,27)	64.9005	0.008350	0.92840	0.06325	0.008601	1.4687
0.5	(12,30)	65.5579	0.007678	0.89686	0.09546	0.010471	1.4924
0.6	(14,34)	65.7585	0.006647	0.85113	0.14222	0.014385	1.5299
0.7	(18,41)	66.1146	0.005342	0.77418	0.22048	0.020694	1.5924
0.8	(21,53)	63.6840	0.003277	0.64082	0.35591	0.035915	1.7098

Table 4: Optimal strategies for constant batch size and varying $E[S_H]$

$E[S_H]$	(m,M)	$\eta_{m,M}$	$\sigma_{m,M}$	$\rho^R_{m,M}$	$\rho^H_{m,M}$	$\zeta_{m,M}$	$\kappa_{m,M}$	Criterion
0.1	(5,19)	42.5629	0.011777	0.97717	0.01105	0.003127	0.005202	1.2055
0.2	(5,20)	42.6349	0.010882	0.96432	0.02479	0.003797	0.006406	1.2098
0.3	(6,21)	43.0088	0.010678	0.94618	0.04315	0.004062	0.007025	1.2169
0.4	(7,23)	43.6622	0.009789	0.92257	0.06764	0.004848	0.008229	1.2291
0.5	(9,25)	44.3229	0.009413	0.88715	0.10344	0.006170	0.010131	1.2496
0.6	(11,29)	45.2483	0.007925	0.83593	0.15614	0.009122	0.014557	1.2843
0.7	(14,36)	46.3561	0.005858	0.74974	0.24441	0.014840	0.023003	1.3468
0.8	(19,48)	45.7194	0.003471	0.58098	0.41555	0.029193	0.044122	1.4764

Proof of Lemma 4.3: We must prove that

$$\sum_{k=0}^{\infty}\sum_{n=0}^{k}\int_0^\infty \sum_{j=0}^{n}\frac{t}{n+1}\sum_{b=j}^{k}(r-b)\beta(\tbinom{*j}{b}|\tbinom{*n}{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)z^k = \tag{8.1}$$
$$-r\hat{G}'(\lambda(1-\beta(z)))-\frac{1}{2}\lambda z\beta'(z)\hat{G}''(\lambda(1-\beta(z))).$$

By rewriting (8.1) in terms of the conditional expectations $E[\sum_{i=1}^{j}B_i|\sum_{i=1}^{n}B_i=k]=j\frac{k}{n}$ we get the result after some simple algebra, as shown below.

$$\sum_{k=0}^{\infty}\sum_{n=0}^{k}\int_0^\infty \frac{t}{n+1}\sum_{j=0}^{n}\left(r-E\left[\sum_{i=1}^{j}B_i\middle|\sum_{i=1}^{n}B_i=k\right]\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)z^k =$$
$$\sum_{k=0}^{\infty}\sum_{n=0}^{k}\int_0^\infty \frac{t}{n+1}(n+1)rz^k\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)+$$
$$-\sum_{k=1}^{\infty}\sum_{n=1}^{k}\int_0^\infty \frac{t}{n+1}\frac{1}{2}n(n+1)\frac{k}{n}z^k\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$r\int_0^\infty t\sum_{n=0}^{\infty}\sum_{k=n}^{\infty}\beta_k^{*n}z^ke^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)-\frac{1}{2}\int_0^\infty\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}tkz^k\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$r\int_0^\infty t\sum_{n=0}^{\infty}e^{-\lambda t}\frac{(\lambda t\beta(z))^n}{n!}\,dG(t)-\frac{1}{2}z\int_0^\infty t\sum_{n=1}^{\infty}n\beta(z)^{n-1}\beta'(z)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$r\int_0^\infty te^{-\lambda t(1-\beta(z))}\,dG(t)-\frac{1}{2}\lambda z\beta'(z)\int_0^\infty t^2e^{-\lambda t(1-\beta(z))}\,dG(t)=$$

$$-r\hat{G}'(\lambda(1-\beta(z))) - \frac{1}{2}\lambda z\beta'(z)\hat{G}''(\lambda(1-\beta(z))).$$

□

As a final remark, the coefficients $\gamma_k(r)$ can be calculated for all $r \geq 1$ by only two applications of the Fast Fourier Transform method, noting that

$$\gamma_k(r) = -r\epsilon_k - \frac{1}{2}\lambda\nu_k,$$

where the ϵ_k and the ν_k are the respective coefficients of the generating functions

$$\mathcal{E}(z) = \sum_{k=0}^{\infty} \epsilon_k z^k := \hat{G}'(\lambda(1-\beta(z))) \quad \text{and} \quad \mathcal{N}(z) = \sum_{k=0}^{\infty} \nu_k z^k := z\beta'(z)\hat{G}''(\lambda(1-\beta(z))).$$

The proof of Lemma 4.4 is an immediate consequence of the following two results, which we will prove separately. Using the notation of Theorem 4.2 we have for $r \geq 1$,

Proposition 1

$$\delta_r = \sum_{k=r}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{r-1}\beta(^{*j}_{b}|^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Proposition 2

$$\theta_{r-1} = \sum_{k=r}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{r-1}b\beta(^{*j}_{b}|^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Proof of Proposition 1: We must show that

$$\sum_{r=1}^{\infty}\sum_{k=r}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{r-1}\beta(^{*j}_{b}|^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)z^r = \tag{8.2}$$
$$\frac{z}{(1-z)(1-\beta(z))}\left\{\frac{1}{\lambda}\left(1-\hat{G}(\lambda(1-\beta(z)))\right) + (1-\beta(z))\hat{G}'(\lambda(1-\beta(z)))\right\}.$$

Interchanging summations, we can rewrite (8.2) as

$$\sum_{k=1}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{k-1}\sum_{r=b+1}^{k} z^r\beta(^{*j}_{b}|^{*n}_{k})\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t). \tag{8.3}$$

Now, writing

$$\sum_{r=b+1}^{k} z^r = \frac{z^{b+1}-z^{k+1}}{1-z},$$

and subsequently,

$$\sum_{b=j}^{k} z^b\beta(^{*j}_{b}|^{*n}_{k}) = E\left[z^{\sum_{i=1}^{j}B_i}\,\middle|\,\sum_{i=1}^{n}B_i = k\right],$$

equation (8.3) becomes

$$\frac{z}{1-z}\sum_{k=1}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\left(E\left[z^{\sum_{i=1}^{j}B_i}\,\middle|\,\sum_{i=1}^{n}B_i = k\right] - z^k\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t),$$

which, after interchanging summations, deconditioning and using the generating functions

$$E\left[z^{\sum_{i=1}^{j}B_i}\right] = \beta(z)^j,$$

leads to

$$\frac{z}{1-z}\sum_{n=1}^{\infty}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\left(\beta(z)^j-\beta(z)^n\right)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$\frac{z}{1-z}\int_0^{\infty}\sum_{n=1}^{\infty}\frac{t}{n+1}\left(\frac{1-\beta(z)^{n+1}}{1-\beta(z)}-(n+1)\beta(z)^n\right)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Now, summation over n and rearranging terms gives the result after a few lines of algebra. □

Proof of Proposition 2: We must show that

$$\sum_{r=1}^{\infty}\sum_{k=r}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{r-1}b\beta\left({}^{*j}_{b}\middle|{}^{*n}_{k}\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)z^{r-1}= \tag{8.4}$$
$$\frac{z\beta'(z)}{(1-z)(1-\beta(z))^2}$$
$$\times\left\{\frac{1}{\lambda}\left(1-\hat{G}(\lambda(1-\beta(z)))\right)+(1-\beta(z))\hat{G}'(\lambda(1-\beta(z)))-\frac{1}{2}\lambda(1-\beta(z))^2\hat{G}''(\lambda(1-\beta(z)))\right\}.$$

The proof follows the same lines as the proof of Proposition 1. Remark that the only essential difference between (8.4) and (8.2) is a factor b. Rewriting (8.4) several times as before gives

$$\sum_{k=1}^{\infty}\sum_{n=1}^{k}\int_0^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{k-1}\sum_{r=b+1}^{k}bz^{r-1}\beta\left({}^{*j}_{b}\middle|{}^{*n}_{k}\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$\int_0^{\infty}\sum_{n=1}^{\infty}\sum_{k=n}^{\infty}\frac{t}{n+1}\sum_{j=0}^{n}\sum_{b=j}^{k}b\frac{z^b-z^k}{1-z}\beta\left({}^{*j}_{b}\middle|{}^{*n}_{k}\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$\frac{1}{1-z}\int_0^{\infty}\sum_{n=1}^{\infty}\frac{t}{n+1}\sum_{j=1}^{n}\sum_{k=n}^{\infty}\left(E\left[\sum_{i=1}^{j}B_iz^{\sum_{i=1}^{j}B_i}-z^k\sum_{i=1}^{j}B_i\,\middle|\,\sum_{i=1}^{n}B_i=k\right]\right)\beta_k^{*n}e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t)=$$
$$\frac{1}{1-z}\int_0^{\infty}\sum_{n=1}^{\infty}\frac{t}{n+1}\sum_{j=1}^{n}\left(E\left[\sum_{i=1}^{j}B_iz^{\sum_{i=1}^{j}B_i}-\sum_{i=1}^{j}B_iz^{\sum_{i=1}^{n}B_i}\right]\right)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t). \tag{8.5}$$

Now we use the fact that the different batches are independent. Then, applying generating functions again, we have

$$E\left[\sum_{i=1}^{j}B_iz^{\sum_{i=1}^{j}B_i}-\sum_{i=1}^{j}B_iz^{\sum_{i=1}^{n}B_i}\right]=jz\beta(z)^{j-1}\beta'(z)(1-\beta(z)^{n-j}). \tag{8.6}$$

Substituting in (8.5) the left-hand side of (8.6) for the right-hand side gives after summation over j

$$\frac{z\beta'(z)}{1-z}\int_0^{\infty}\sum_{n=1}^{\infty}\frac{t}{n+1}\left(\frac{1-(n+1)\beta(z)^n(1-\beta(z))-\beta(z)^{n+1}}{(1-\beta(z))^2}-\frac{1}{2}n(n+1)\beta(z)^{n-1}\right)e^{-\lambda t}\frac{(\lambda t)^n}{n!}\,dG(t).$$

Finally, summation over n and rearranging terms gives the result. □

Acknowledgement

The authors are indebted to Henk Tijms for his helpful discussions which significantly improved the exposition of this paper.

References

[1] Abate, S. and Whitt, W. (1992). The Fourier-series method for inverting transforms of probability functions. *Queueing Systems* 10: 5–88.

[2] Altiok, T. (1989). (R, r) Production/Inventory Systems. *Operations Research* 37,2:266–276.

[3] Asmussen, S. (1987). *Applied Probability and Queues*, John Wiley and Sons, New York.

[4] De Kok, A.G., Tijms, H.C. and Van der Duyn Schouten, F.A. (1984). Approximations for the single-product production-inventory problem with compound Poisson demand and service level constraints, *Advances of Applied Probability*, 16: 378–301.

[5] De Kok, A.G., Tijms, H.C. and Van der Duyn Schouten, F.A. (1985). Inventory levels to stop and restart a single machine producing one product, *European Journal of Operational Research*, 20: 239–247.

[6] Doshi, B.T., Van der Duyn Schouten, F.A. and Talman, A.J.J. (1978). A production-inventory control model with a mixture of back-orders and lost-sales. *Management Science* 24: 1078–1086.

[7] Dshalalow, J.H. (1997). Queueing Systems with state dependent parameters, in: *Frontiers in Queueing*, edited by J.H. Dshalalow, CRC Press, New York.

[8] Federgruen, A. and Tijms, H.C. (1980). Computation of the Stationary Distribution of the Queue Size in an M/G/1 Queueing System with Variable Service Rate. *Journal of Applied Probability* 17: 515–522.

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

[10] Heyman, D.P. (1968). Optimal Operating Policies for $M/G/1$ Queueing Systems. *Operations Research* 16:362–382.

[11] Holewijn, P.J. and Hordijk, A. (1975). On the Convergence of Moments in Stationary Markov Chains. *Stochastic Processes and their Applications* 3:55–64.

[12] Lee, H.S. (1995). The Optimal (r, S) Policy in a Single Item Production/Inventory System with Setup Times. *Journal of the Operations Research Society of Japan* 38:141–161.

[13] Lee, H.S. and Srinivasan, M.M. (1989). Control Policies for the $M^X/G/1$ Queueing System. *Management Science*, 35:708–721.

[14] Nishimura, S. and Jiang, Y. (1995). An $M/G/1$ Vacation Model with Two Service Modes. *Probability in the Engineering and Informational Sciences* 9:355–374.

[15] Nobel, R.D. (1998). A regenerative approach for an $M^X/G/1$ queue with two service modes. *Automatic Control and Computer Sciences*, 1:3–14 (in Russian).

[16] Nobel, R.D. and Ridder, A.A.N. (1998). Optimal control for an $M^X/G/1/N+1$ queue with two service modes. [submitted]

[17] Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T. (1986). *Numerical Recipes*, Cambridge University Press, Cambridge.

[18] Silver, E. and Peterson, R. (1985). *Decision Systems for Inventory Management and Production Planning*, 2nd ed., John Wiley and Sons, New York.

[19] Srinisavan, M.M. and Lee, H.S. (1991). Random Review Production/Inventory Systems with compound Poisson Demands and Arbitrary Processing Times. *Management Science*, 37:813–833.

[20] Tijms, H.C. (1980). An algorithm for average cost denumerable state semi-Markov decision problems with applications to controlled production and queueing systems, *Recent Developments in Markov Decision Processes*, eds. R. Hartley, L.C. Thomas and D.J. White, pp 143–179, Academic Press, New York.

[21] Tijms, H.C. (1994). *Stochastic Models: An Algorithmic Approach*, John Wiley and Sons, New York.

[22] Yamada, K. and Nishimura, S. (1994). A Queueing system with a Setup Time for Switching of the Service Distribution. *Journal of the Operations Research Society of Japan*, 37:271–286.

Received: 3/10/1999
Revised: 6/4/2000
Accepted: 8/1/2000