Available online at www.sciencedirect.com

SCIENCE DIRECT

Int. J. Production Economics 88 (2004) 105–119

international journal of production economics

www.elsevier.com/locate/dsw

The stochastic dynamic production/inventory lot-sizing problem with service-level constraints

S. Armagan Tarim[a], Brian G. Kingsman[b,*]

[a] Department of Computer Science, University of York, York YO10 5DD, UK

[b] Department of Management Science, The Management School, Lancaster University, Lancaster LA1 4YX, UK

Received 17 April 2003; accepted 26 May 2003

Abstract

This paper addresses the multi-period single-item inventory lot-sizing problem with stochastic demands under the "static–dynamic uncertainty" strategy of Bookbinder and Tan (Manage. Sci. 34 (1988) 1096). In the static-dynamic uncertainty strategy, the replenishment periods are fixed at the beginning of the planning horizon, but the actual orders are determined only at those replenishment periods and will depend upon the demand that is realised. Their solution heuristic was a two-stage process of firstly fixing the replenishment periods and then secondly determining what adjustments should be made to the planned orders as demand was realised. We present a mixed integer programming formulation that determines both in a single step giving the optimal solution for the "static–dynamic uncertainty" strategy. The total expected inventory holding, ordering and direct item costs during the planning horizon are minimised under the constraint that the probability that the closing inventory in each time period will not be negative is set to at least a certain value. This formulation includes the effect of a unit variable purchase/production cost, which was excluded by the two-stage Bookbinder–Tan heuristic. An evaluation of the accuracy of the heuristic against the optimal solution for the case of a zero unit purchase/production cost is made for a wide variety of demand patterns, coefficients of demand variability and relative holding cost to ordering cost ratios. The practical constraint of non-negative orders and the existence of the unit variable cost mean that the replenishment cycles cannot be treated independently and so the problem cannot be solved as a stochastic form of the Wagner–Whitin problem, applying the shortest route algorithm.

Keywords: Inventory; Chance-constrained programming; Integer programming; Modelling; Optimisation

1. Introduction

The study of lot-sizing began with Wagner and Whitin (1958), and there is now a sizeable literature in this area extending the basic model to consider capacity constraints, multiple items, multiple stages, etc. However, most previous work on lot-sizing has been directed towards the

*Corresponding author. Tel.: 44-1524-593-848; fax: 44-1524-844-885.

E-mail addresses: at@cs.york.ac.uk (S.A. Tarim), b.kingsman@lancaster.ac.uk, b.kingsman@uk.ac.lancaster (B.G. Kingsman).

0925-5273/03/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0925-5273(03)00182-8

deterministic case. The reader is directed to De Bodt et al. (1984), Potts and Van Wassenhove (1992), Kuik et al. (1994) and Kimms (1997) for a review of lot-sizing techniques.

The practical problem is that in general much, if not all, of the future demands have to be forecast. Point forecasts are typically treated as deterministic demands. However, the existence of forecast errors radically affects the behaviour of the lot-sizing procedures based on assuming the deterministic demand situation. Forecasting errors lead both to stockouts occurring with unsatisfied demands and to larger inventories being carried than planned. The introduction of safety stocks in turn generates even larger inventories and also more orders. It is reported by Davis (1993) that a study at Hewlett-Packard revealed the fact that 60% of the inventory investment in their manufacturing and distribution system is due to demand uncertainty.

There has been increasing recognition as illustrated by Wemmerlov (1989) that future lot-sizing studies need to be undertaken on stochastic and dynamic environments that have at least a modicum of resemblance to reality. Inevitably, the forecast errors have to be taken into account in planning the future lot-sizes. Similar concerns have been expressed by Silver: "One should not necessarily use a deterministic lot-sizing rule when significant uncertainty exists. A more appropriate strategy might be some form of probabilistic modelling."

Silver (1978) suggested a heuristic procedure for the stochastic lot-sizing problem assuming that the forecast errors are normally distributed. A similar heuristic, having a different objective function, was presented by Askin (1981). Bookbinder and Tan (1988) proposed another heuristic, under the "static–dynamic uncertainty" strategy. In this strategy, the replenishment periods are fixed at the beginning of the planning horizon and the actual orders at future replenishment periods are determined only at those replenishment periods depending upon the realised demand. The total expected cost is minimised under the minimal service-level constraint. In this paper, we propose a mixed integer programming formulation to solve the stochastic dynamic lot-sizing problem to optimality under the "static–dynamic uncertainty" strategy of Bookbinder and Tan.

The optimal solution to the problem is the (s, S) policy with different values for each period in the time varying demand situation. Where the demand level changes slowly, it is usually satisfactory to use a steady-state analysis with constant S and s values updated once per year. However, this is inappropriate if the average demand can change significantly from period to period. This presents a non-stationary problem where the two control parameters change from period to period. The uncertainty in the timing of future replenishments caused by an (s, S) policy may be unattractive from an operational standpoint.

Although the inventory problems with stationary demand assumption are well known and extensively studied, very little has appeared on the non-stationary stochastic demand case. Recently, Sox (1997) and Martel et al. (1995) have described static control policies under the non-stationary stochastic demand assumption in a rolling horizon framework. Sox (1997) presents a mixed integer non-linear formulation of the dynamic lot-sizing problem with dynamic costs, and develops a solution algorithm that resembles the Wagner–Whitin algorithm but with some additional feasibility constraints. Martel et al. (1995) transform the multiple item procurement problem into a multi-period static decision problem under risk. Other notable works on non-stationary stochastic demand adopt (s, S) or base-stock policies and are due to Iida (1999), Sobel and Zhang (2001), and Gavirneni and Tayur (2001). In Iida (1999), the periodic review dynamic inventory problem is considered and it is shown that near myopic policies are sufficiently close to optimal decisions for the infinite horizon inventory problem. In Sobel and Zhang (2001), it is assumed that demand arrives simultaneously from a deterministic source and a random source, and proven that a modified (s, S) policy is optimal under general conditions. Gavirneni and Tayur (2001) use the derivative of the cost function that can handle a much wider variety of fluctuations in the problem parameters. The above studies have adopted either static control policies in a rolling horizon framework or dynamic control policies like (s, S), although the static-dynamic uncertainty model is a more accurate representation of industrial practice as pointed out by Sox (1997).

Most companies use MRP in some form for their production planning and thus ordering on suppliers. They typically issue advance schedules of requirements. In talking to companies in the supply chain who are in this situation a common complaint is that their customers continually change their schedules, the timing of orders as well as the size of orders. It is the changing of the timing that they find worse. This issue of system nervousness is an active current research area. If there is to be more co-operation and co-ordination in supply chains, then a model that attempts to determine a schedule for the timing of orders in advance taking account of the stochastic demand, which remains fixed, is a contribution of practical interest. This need to fix the deliveries in advance, whilst allowing reasonable flexibility in the order size has been at the heart of the problem of buying raw materials on fluctuating price markets. You have to determine which future months or half months in which to have your delivery of wheat or cocoa, etc. Once this is fixed, suppliers do not allow it to be changed, although the quantities ordered can be varied to some extent. Waiting until the month or half month of delivery and using the standard (s, S) model to decide whether or not to place an order and then its size is not a good option, as the price you pay tends to be much higher as sellers know you are desperate for the material. Thus, investigating models and solution procedures for the static-dynamic uncertainty strategy is potentially important from the practical application perspective.

2. Problem statement

The stochastic dynamic lot-sizing problem can be formulated either by assuming a penalty cost for each stockout and unsatisfied demand or by minimising the ordering and inventory costs subject to satisfying some customer service-level criterion. In this paper we consider meeting a specified customer service case, where we use a "minimal service-level" criterion. A detailed investigation of the minimal service-level criterion is given by Chen and Krass (2001). For a discussion of different service constraints see Van Houtum and Zijm (2000) and Rosling (1999). Remember that demand is not deterministic with different values in each period. Thus, the service-level criterion in this case is defined as specifying a minimum probability, say α, that at the end of every period the net inventory will not be negative. In general, finding the solution to this problem can be formulated as solving a chance-constrained programming model (see for example Birge and Louveaux, 1997). It can be expressed as the minimisation of the total expected cost, $E\{\mathrm{TC}\}$, over the N-period planning horizon subject to the service-level constraints, as given below:

$$\text{minimise } E\{\mathrm{TC}\} = \int_{d_1}\int_{d_2}\cdots\int_{d_N}\sum_{t=1}^{N}(a\delta_t + hI_t + vX_t\} \times g_1(d_1)g_2(d_2)\ldots g_N(d_N)\mathrm{d}(d_1)\mathrm{d}(d_2)\ldots\mathrm{d}(d_N) \quad (1)$$

subject to

$$\delta_t = \begin{cases} 1 & \text{if } X_t > 0, \\ 0 & \text{otherwise,} \end{cases} \quad t = 1, \ldots, N, \quad (2)$$

$$I_t = I_0 \sum_{i=1}^{t}(X_i - d_i), \quad t = 1, \ldots, N, \quad (3)$$

$$\Pr\{I_t \geqslant 0\} \geqslant \alpha, \quad t = 1, \ldots, N, \quad (4)$$

$$X_t, I_t \geqslant 0, \quad t = 1, \ldots, N,$$

where N is the number of periods in the planning horizon, X_t the replenishment order placed and received (i.e., no lead-time) in period t, δ_t a {0,1} variable that takes the value of 1 if a replenishment order is placed in period t and 0 otherwise, I_t the inventory level at the end of period t, I_0 the stock on hand at the beginning of period 1, d_t the demand in period t, a the fixed procurement cost, v the marginal cost of purchasing an item, h the linear holding cost incurred on any unit carried in inventory over from one period to the next.

The demand d_t in period t is considered as a random variable with known probability density function, $g_t(d_t)$. The distribution of demand may vary from period to period. Demands in different time periods are assumed independent. A fixed procurement (ordering or set-up) cost, a, is incurred each time a replenishment order is placed, whatever the size of the order. In addition to the fixed ordering cost, a variable purchasing cost is incurred depending on the size of the order, vX_t. A replenishment order is assumed to arrive instantaneously at the beginning of each period, before the demand in that period occurs. A linear holding cost h is incurred on any unit carried in inventory over from one period to the next. For simplicity in the analysis, inventory holding costs are assumed to be incurred only on the inventory at the end of each period. The extension to taking the average inventory over a period is straightforward. The probability that at the end of each and every time period the net inventory will not be negative is set to be at least α. Hence, it is implicitly assumed that, since normally the desired service level is quite high, the value α incorporates the perception of the cost of backorders, so that shortage cost can be ignored in the model. The above assumptions are valid for the rest of the paper.

3. The Bookbinder–Tan strategy for the probabilistic lot-sizing problem

One of the few attempts to produce more implementable near optimal solutions to the stochastic dynamic lot-size problem is given by Bookbinder and Tan (1988). In Bookbinder and Tan's paper three strategies were analysed, named "static uncertainty", "dynamic uncertainty" and "static–dynamic uncertainty". In all cases, their analysis was based on a service-level constraint on the probability of a stockout at the end of each period.

The "static uncertainty" decision rule assumes that the values of all the decision variables, the timing and size of replenishment orders, are determined at the beginning of the planning horizon. All the X_t can be considered as deterministic decision variables. The values and timing of the replenishment orders is determined by initially reducing the problem to the equivalent deterministic lot-sizing model given below:

$$\text{minimise } E\{\text{TC}\} = \sum_{t=1}^{N} (a\delta_t + hE\{I_t\} + vX_t) \quad (5)$$

subject to

$$\delta_t = \begin{cases} 1 & \text{if } X_t > 0, \\ 0 & \text{otherwise,} \end{cases} \quad t = 1, \ldots, N, \quad (6)$$

$$E\{I_t\} = I_0 + \sum_{i=1}^{t} X_i - \sum_{i=1}^{t} E\{d_i\}, \quad t = 1, \ldots, N, \quad (7)$$

$$\sum_{i=1}^{t} X_i \geqslant G_{D(t)}^{-1}(\alpha) - I_0, \quad t = 1, \ldots, N, \quad (8)$$

$$X_t, E\{I_t\} \geqslant 0, \quad t = 1, \ldots, N.$$

The above formulation is solved by the Wagner–Whitin algorithm. Bookbinder and Tan give the equivalent linear constraint, $\sum_{i=1}^{t} X_i \geqslant G_{D(t)}^{-1}(\alpha) - I_0$, for the service-level constraint $Pr\{I_t \geq 0\} \geqslant \alpha$, where $G_{D(t)} = G_{d_1+d_2+\cdots+d_t}(.)$ is the cumulative probability distribution function of $d_1 + d_2 + \cdots + d_t$. It is assumed that G is strictly increasing, therefore G^{-1} is uniquely defined. They also examine the analogy between the deterministic model and the "static uncertainty" model, and conclude that there is correspondence between the respective terms of the two models.

Since $G_{d_1}^{-1}(\alpha) \leqslant G_{d_1+d_2}^{-1}(\alpha) \leqslant \cdots \leqslant G_{D(N)}^{-1}(\alpha)$, the constraint $\sum_{i=1}^{t} X_i \geqslant G_{D(t)}^{-1}(\alpha) - I_0$ is binding for $t = N$ and gives $\sum_{t=1}^{N} X_t = G_{D(N)}^{-1}(\alpha) - I_0$, which implies that the total direct item cost, $\sum_{t=1}^{N} vX_t$, is constant. Therefore, Bookbinder and Tan ignore the total direct item cost term, which has no effect on determination of the best schedule, as in the deterministic problem.

Another approach for determining the timing and size of the replenishment orders is the "dynamic uncertainty" strategy. In the "dynamic uncertainty" strategy decisions are made every period on the basis of the demands that have become known as you move forward through time. This approach ignores the ordering cost a, and may require a replenishment order almost every period. Obviously, for large ratios of a/h this result is clearly undesirable.

Bookbinder and Tan then combined the features of both strategies to give a two-stage heuristic, which they called the "static–dynamic uncertainty" approach. The first stage was to fix the timing of the replenishment orders using the "static uncertainty" model. Exploiting the analogy between the deterministic model and the "static uncertainty" model, the replenishment periods can be determined by using dynamic programming (Wagner and Whitin, 1958), network flows (Zangwill, 1968) or integer programming (Karni, 1981). The second stage was then to calculate the adjustments made to those orders at the times they were scheduled as the demand was realised over the planning horizon. These adjustments were expressed as margins to add to the total demand received over all the periods since the immediately preceding order was received.

The Bookbinder and Tan analysis for the second stage is briefly as follows. Consider a review schedule, which has m reviews over the N period planning horizon with orders arriving at $\{T_1, T_2, \ldots, T_m\}$, where $T_j > T_{j-1}$, $T_m \leqslant N$. Remember that the T_i have been determined at the first stage. For convenience $T_1 = 1$ is defined as the start of the planning horizon and $T_{m+1} = N + 1$ the period immediately after the end of the horizon. The review schedule may be generalised to consider the case where $T_1 > 1$, if the opening stock, I_0, is sufficient to cover the immediate needs at the start of the planning horizon. The associated stock reviews will take place at the beginning of periods T_i, $i = 1, \ldots, m$. In the considered dynamic review and replenishment policy, clearly the orders X_t are all equal to zero except at the replenishment periods $T_1, T_2, \ldots, T_m$. The inventory level I_t carried from period t to period $t + 1$ is the opening stock plus any orders that have arrived up to and including period t less the total demand to date. Hence it is given as

$$I_t = I_0 + \sum_{j=1}^{i} X_{T_j} - \sum_{k=1}^{t} d_k, \quad T_i \leqslant t < T_{i+1}, \quad i = 1, \ldots, m. \tag{9}$$

Bookbinder and Tan transform the decision variable X_{T_i} to a new variable $Y_{T_i} \in R$, which may be interpreted as the change from period T_{i-1} to period T_i in the desired buffer stock level and is given as follows:

$$\begin{aligned} X_{T_1} &= Y_{T_1}, \\ X_{T_2} &= Y_{T_2} + d_{T_1} + \cdots + d_{T_2 - 1}, \\ &\vdots \\ X_{T_m} &= Y_{T_m} + d_{T_{m-1}} + \cdots + d_{T_m - 1}. \end{aligned} \tag{10}$$

Hence, the replenishment size X_{T_i} can be determined by means of the deterministic decision variable Y_{T_i} after the demand $\sum_{k=T_{i-1}}^{T_i - 1} d_k$ is realised. It follows from Eq. (10) that

$$\sum_{j=1}^{i} X_{T_j} = \sum_{j=1}^{i} Y_{T_j} + \sum_{k=1}^{T_i - 1} d_k, \quad i = 1, \ldots, m, \tag{11}$$

and that the inventory level I_t at the end of period t, given by Eq. (9), can be expressed as

$$I_t = I_0 + \sum_{j=1}^{i} Y_{T_j} - \sum_{k=T_i}^{t} d_k, \quad T_i \leqslant t < T_{i+1}, \quad i = 1, \ldots, m. \tag{12}$$

Therefore, the random components of the closing inventory level, I_t, depends only on the demand since the most recent stock review in period T_i, $i = 1, \ldots, m$. Note that Y_{T_i} is not a random variable; rather, it is a deterministic decision variable whose value is determined using a linear programming model.

As in the static model, Bookbinder and Tan say that the probabilistic constraints,

$$\sum_{i=1}^{j} Y_{t_i} \geqslant G^{-1}_{d_{t_j} + d_{t_j+1} + d_{t_j+2} + \cdots + d_{t_{j+1}-1}}(\alpha) - I_0, \quad j = 1, 2, \ldots, m,$$

hold as an equality at optimality. By substituting $G^{-1}_{d_{t_m} + d_{t_m+1} + d_{t_m+2} + \cdots + d_{t_{m+1}-1}}(\alpha) - I_0$ for $\sum_{i=1}^{m} Y_{t_i}$, we observe that their total unit variable cost expression, $v\left\{\sum_{i=1}^{m} Y_{t_i} + \sum_{i=1}^{t_m - 1} E\{d_i\}\right\}$, takes a constant value. In other words, we can conclude that the unit variable cost has no role in replenishment policies determined according to the Bookbinder–Tan heuristic.

4. An optimisation model for the probabilistic lot-sizing problem

The Bookbinder and Tan static-dynamic procedure splits the solution into two independent stages, ignoring the interactions between them. It is therefore a heuristic solution procedure. The question then arises as to its accuracy and the likely level of cost penalty that it incurs.

To evaluate the accuracy of the heuristic, the problem must be formulated in a way that gives the optimal solution under the "static–dynamic uncertainty" strategy. This will allow the simultaneous determination of the number and timing of the replenishments and the information necessary to determine the size of the replenishment orders that will minimise the expected costs of meeting demand over some finite planning horizon, given a set of forecasts of the demands and a service-level constraint on the probability of a stockout. This is done in this section. It basically models the Bookbinder and Tan transformation process in a way that allows the simultaneous determination of the timing and size of the orders as a single step. The initial decisions depend upon what information can be expected to become available in the future and how best to react to this information. Therefore, the replenishment times to use are determined at the beginning of the planning horizon considering the interdependency between the stock levels to have available at the start of those periods and their timing. The actual order quantities at future replenishment periods are determined only at those replenishment periods and will depend upon the demand that is realised period by period over the planning horizon. It is assumed that negative orders are not allowed, so that if the actual stock exceeds the order-up-to-level for that review, this excess stock will be carried forward and not returned to the supply source.

4.1. Reformulating the constraints of the chance-constrained model

We start with Eq. (12) of Section 3, but with the difference that the number and timing of the reviews are still variables to be determined and not preset values.

Defining $R_{T_i} = I_0 + \sum_{j=1}^{i} Y_{T_j}$ and substituting for I_t, Eq. (12) becomes

$$I_t = R_{T_i} - \sum_{k=T_i}^{t} d_k, \qquad T_i \leqslant t < T_{i+1}, \quad i = 1, \ldots, m. \tag{13}$$

Note that R_{T_i} may be interpreted as an order-up-to-level to which stock should be raised after receiving an order at the ith review period T_i, and $R_{T_i} - \sum_{k=T_i}^{t} d_k$ is the end of period inventory. Thus, instead of working in terms of decision variables Y_{T_i}, as in the model proposed by Bookbinder and Tan, the problem can be expressed in terms of these new decision variables R_{T_i}. The problem is to determine the number of reviews, m, the T_i, and the associated R_{T_i} for $i = 1, \ldots, m$.

If there is no replenishment scheduled for period t, then R_t equals the opening inventory level in period t. Now, Eq. (13) can be expressed more simply as

$$I_t = R_t - d_t, \quad t = 1, \ldots, N. \tag{14}$$

It follows that the variable R_t must be equal to I_{t-1} if no order is received in period t and equal to the order-up-to-level if there is a review and the receipt of an order. The first case applies if no stock review takes place in period t, which is indicated by the integer variable $\delta_t = 0$. If $\delta_t = 0$ then R_t must equal I_{t-1}. This is achieved by the two linear inequalities given by Eq. (15), since the constraints become $R_t \leqslant I_{t-1}$ and $R_t \geqslant I_{t-1}$, respectively, for $\delta_t = 0$:

$$\begin{aligned} &R_t - I_{t-1} \leqslant M\delta_t, \\ &R_t \geqslant I_{t-1}, \end{aligned} \qquad t = 1, \ldots, N, \tag{15}$$

where M is some very large positive number.

Whilst if $\delta_t = 1$ then the constraints require R_t to lie between infinity and I_{t-1} satisfying the other condition on R_t. The values for the order-up-to-level variables, R_t, when $\delta_t = 1$ are then those that give the minimum expected costs $E\{\mathrm{TC}\}$. The desired opening stock levels, as required for the solution to the problem, will then be those values of R_t for which $\delta_t = 1$.

It is, therefore, clear from the above explanations that constraints (6) and (7) can be replaced with Eqs. (15) and (14), respectively.

As mentioned before, α is the desired minimum probability that the net inventory level in any time period will actually be non-negative and determined subjectively. With this regard the chance constraint

$$\Pr\{I_t \geqslant 0\} \geqslant \alpha, \quad t = 1, \ldots, N \tag{16}$$

can, using Eq. (13), be written alternatively as

$$\Pr\left\{ R_{T_i} \geqslant \sum_{k=T_i}^{t} d_k \right\} \geqslant \alpha, \quad t = 1, \ldots, N, \tag{17}$$

which implies

$$G_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(R_{T_i}) \geqslant \alpha, \quad T_i \leqslant t < T_{i+1}, \quad i = 1, \ldots, m, \tag{18}$$

and

$$I_t \geqslant G^{-1}_{d_{T_i}+d_{T_{i+1}}+\ldots+d_t}(\alpha) - \sum_{k=T_i}^{t} d_k, \quad T_i \leqslant t < T_{i+1}, \quad i = 1, \ldots, m. \tag{19}$$

The right-hand side of Eq. (19) can be calculated or possibly read from a table, once the form of $g_t(.)$ is selected.

In Eq. (19), $G^{-1}_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(.)$ can only be determined after the replenishment periods T_i have been determined. But, as these are chosen to minimise the expected costs, the stock replenishment periods cannot be determined until the appropriate $G^{-1}_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(.)$ values to use in the model are known. There is an obvious circularity here in trying to solve the problem. Since Bookbinder and Tan separate the determination of the timing of the replenishment orders and the adjustments to those orders, they avoid the circularity by sacrificing optimality. One way to overcome this problem, by not sacrificing optimality, is to formulate it as a mixed integer linear programming model.

Since the problem has a finite planning horizon of N periods, $G^{-1}_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(.)$ can be calculated for all relevant cases. If the 0/1 integer variable P_{tj} is defined as taking a value of 1 if the most recent order prior to period t was in period $t-j+1$ and zero elsewhere, then $G^{-1}_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(\alpha)$ can be expressed as

$$G^{-1}_{d_{T_i}+d_{T_{i+1}}+\cdots+d_t}(\alpha) = \sum_{j=1}^{t} G^{-1}_{d_{t-j+1}+d_{t-j+2}+\cdots+d_t}(\alpha) P_{tj}, \quad t = 1, \ldots, N,$$

and similarly Eq. (19) can be expressed as

$$I_t \geqslant \sum_{j=1}^{t} \left(G^{-1}_{d_{t-j+1}+d_{t-j+2}+\cdots+d_t}(\alpha) - \sum_{k=t-j+1}^{t} d_k \right) P_{tj}, \quad t = 1, \ldots, N. \tag{20}$$

The result $P_{tt} = 1$ means that the stock review was in period 1, the start of the planning horizon, whilst $P_{t1} = 1$ means that the stock review was at the start of period t itself. There can at most be only one most recent order received prior to period t. Thus, the P_{tj} must satisfy

$$\sum_{j=1}^{t} P_{tj} = 1, \quad t = 1, \ldots, N. \tag{21}$$

Three other conditions as given below are necessary to identify uniquely the period in which the most recent review prior to any period t took place.

- If $\delta_{t-j+1} = 1$ and $\sum_{k=t-j+2}^{t} \delta_k = 0$, so all subsequent δ_k for $k = t-j+2, t-j+3, \ldots, t$ are 0, then we must have $P_{tj} = 1$, as in these circumstances period $t-j+1$ had the most recent stock review prior to period t.
- If $\delta_{t-j+1} = 0$ and $\sum_{k=t-j+2}^{t} \delta_k = 0$, then $P_{tj} = 0$ since the most recent review prior to period t must have been earlier than $t-j+1$.
- If $\delta_{t-j+1} = 1$ and $\sum_{k=t-j+2}^{t} \delta_k \geqslant 1$, then $P_{tj} = 0$, since other reviews prior to period t occur after period $t+j-1$.

All of these three conditions can be satisfied by the equality condition given in Eq. (21) and the single constraint given below, which are designed to identify uniquely the periods in which the most recent order prior to t took place: 两个条件足够确保3种情况的满足

$$P_{tj} \geqslant \delta_{t-j+1} - \sum_{k=t-j+2}^{t} \delta_k, \quad t = 1, \ldots, N, \quad j = 1, \ldots, t. \tag{22}$$

4.2. The mixed integer programming model to determine the optimal policy under the static dynamic uncertainty strategy

The chance-constrained programming model can be expressed as minimising the objective function given earlier as Eq. (1) subject to the constraints given by Eqs. (14), (15), (20)–(22) and the non-negativity conditions and the 0/1 integer values for δ_t and P_{tj}. Since the decision rule for the above stochastic optimisation problem is chosen to be the "static–dynamic uncertainty" strategy, the timing of the replenishments, obtained from δ_t and P_{tj}, must be decided once and for all, before any of the demands, d_t, become known. Therefore, the expectation operator must be applied to the stochastic variables I_t, R_t and d_t in the constraint equations and objective function. In the chance-constrained programming model given below, the expected value of I_t and R_t are denoted by $E\{I_t\}$ and $E\{R_t\}$, respectively.

Such an analysis is completed at the beginning of the horizon by taking expectations. Hence the deterministic equivalent model for the chance-constrained programming model under the static-dynamic uncertainty strategy may be obtained by taking expectations (see Bookbinder and Tan). The model then is

$$\text{minimise } E\{\text{TC}\} = \sum_{t=1}^{N}(a\delta_t + hE\{I_t\} + vE\{R_t\} - vE\{I_{t-1}\}), \quad (23)$$

where the total expected direct item cost $\sum_{t=1}^{N} vE\{X_t\}$ is written as $\sum_{t=1}^{N}(vE\{R_t\} - vE\{I_{t-1}\})$ following the definition of order-up-to-levels, R_t,

subject to

$$E\{I_t\} = E\{R_t\} - E\{d_t\}, \quad t = 1, \ldots, N, \quad (24)$$

$$E\{R_t\} \geqslant E\{I_{t-1}\}, \quad t = 1, \ldots, N, \quad (25)$$

$$E\{R_t\} - E\{I_{t-1}\} \leqslant M\delta_t, \quad t = 1, \ldots, N, \quad (26)$$

$$E\{I_t\} \geqslant \sum_{j=1}^{t}\left(G^{-1}_{d_{t-j+1}+d_{t-j+2}+\cdots+d_t}(\alpha) - \sum_{k=t-j+1}^{t} E\{d_k\}\right)P_{tj}, \quad t = 1, \ldots, N, \quad (27)$$

$$\sum_{j=1}^{t} P_{tj} = 1, \quad t = 1, \ldots, N, \quad (28)$$

$$P_{tj} \geqslant \delta_{t-j+1} - \sum_{k=t-j+2}^{t} \delta_k, \quad t = 1, \ldots, N, \quad j = 1, \ldots, t, \quad (29)$$

$$E\{I_t\}, E\{R_t\} \geqslant 0, \quad \delta_t, P_{tj} \in \{0, 1\}, \quad t = 1, \ldots, N, \quad j = 1, \ldots, t.$$

This model thus determines the optimum number of replenishments as well as the optimum replenishment schedule, the timing of the replenishments, together with the optimum values to use for dynamically determining the sizes of the replenishment orders as demand is realised, that give the minimum expected total costs. The problem is to determine the values of the 0/1 integer variables, δ_t for $t = 1, \ldots, N$ and P_{tj} for $j = 1, \ldots, t$, $t = 1, \ldots, N$, and the non-negative continuous variables $E\{I_t\}$ and $E\{R_t\}$ for $t = 1, \ldots, N$, that minimise the objective function. To comply with the non-negativity constraint on the lot-sizes of the original model, we must have $E\{R_t\} - E\{I_{t-1}\} \geqslant 0$, $t = 1, \ldots, N$ which is already given in Eq. (15) and incorporated into the model. The times of the stock reviews are given by the values of i such that $\delta_i = 1$. The associated order-up-to-levels, for each i, are given as $E\{R_i\}$.

In Section 3, it was shown that the Bookbinder–Tan heuristic ignores the unit variable cost in determination of the best schedule as in the deterministic problem. However, in contrast to the Bookbinder–Tan heuristic, the above formulation does not ignore the unit variable cost. A rearrangement of the unit variable cost terms in the objective function gives

$$\begin{aligned} v\sum_{t=1}^{N}(E\{R_t\} - E\{I_{t-1}\}) &= -vI_0 + v\sum_{t=1}^{N-1}(E\{R_t\} - E\{I_t\}) + vE\{R_N\} \\ &= -vI_0 + v\sum_{t=1}^{N-1}(E\{d_t\}) + vE\{R_N\} \end{aligned}$$

using Eq. (24) for $t = 1, \ldots, N-1$. Thus, one can write the total unit variable cost component as $\phi + vE\{R_N\}$, where ϕ is a constant with a value equal to v multiplied by the given opening stock plus the expected demand over the first $N-1$ periods. It is clear that the total unit variable cost is a function of the model variable $E\{R_N\}$, which is a variable rather than a constant. Thus the solution will be affected by the variable production cost.

5. A numerical example

To illustrate the technique we shall use the demand forecasts presented in Table 1. The initial inventory level is taken as zero. It is assumed that the demand in each period is normally distributed about the forecast value with a constant coefficient of variation, $C = \sigma_t/\mu_t = 0.333$. The other parameters of the problem are $a = \$2500$ per order, $h = \$1$ per unit per period, and $\alpha = 0.95$ ($z_{\alpha=0.95} = 1.645$). Since BT heuristic ignores the unit variable cost, the example takes into account only the holding and ordering costs (i.e., $v = 0$). In Table 2, the calculated values of the coefficient $G^{-1}_{d_{t-j+1}+d_{t-j+2}+\cdots+d_t}(\alpha)$ are given. For instance, element ($t = 7$, $j = 3$) gives the value of $G^{-1}_{d_5+d_6+d_7}(0.95) = 2833$ which corresponds to the opening inventory level in period $t - j + 1 = 5$ to satisfy the demands of periods 5–7 with probability of at least $\alpha = 95\%$. These values can be calculated using any spreadsheet program. The solution to the problem is calculated using both BT heuristic and the optimisation model derived in the previous section, denoted by (TK).

The first step in the two-step BT heuristic is the determination of the replenishment times. The replenishment times are fixed by means of the "static uncertainty" model, applying the Wagner–Whitin algorithm to the problem formulated as a deterministic one using the analogy of Bookbinder and Tan. This gives periods 1, 5 and 7 as the replenishment periods. Following that the "static–dynamic uncertainty" model is used to determine the order-up-to-levels for these replenishment periods. The results are given in Table 3 and Fig. 1.

The same problem is also solved by means of the deterministic equivalent mixed integer

Table 1
Forecasts of period demands

Period (k)	1	2	3	4	5	6	7	8	9	10
$E\{d_k\}$	800	850	700	200	800	700	650	600	500	200

Table 2
Calculated $G^{-1}_{d_{t-j+1}+d_{t-j+2}+\cdots+d_t}(\alpha) = \sum_{k=t-j+1}^{t} E\{d_k\} + z_{0.95}C\left(\sum_{k=t-j+1}^{t} E^2\{d_k\}\right)^{1/2}$ values

t	j									
	1	2	3	4	5	6	7	8	9	10
1	1239									
2	1316	2290								
3	1084	2154	3096							
4	310	1299	2364	3304						
5	1239	1452	2293	3304	4223					
6	1084	2083	2293	3106	4096	5003				
7	1006	1874	2833	3042	3841	4818	5718			
8	929	1735	2568	3508	3716	4507	5475	6370		
9	774	1528	2307	3127	4056	4264	5050	6013	6904	
10	310	995	1742	2518	3335	4264	4471	5256	6219	7110

Table 3
Replenishment policy of BT

Period (t)	1	2	3	4	5	6	7	8	9	10
Order-up-to-level ($E\{R_t\}_{\delta_t=1}$)	3304	—	—	—	2083	—	2518	—	—	—
Exp. opening inv. ($E\{R_t\}$)	3304	2504	1654	954	2083	1283	2518	1868	1268	768
Exp. closing inv. ($E\{I_t\}$)	2504	1654	954	754	1283	583	1868	1268	768	568
Total expected cost	19,704									

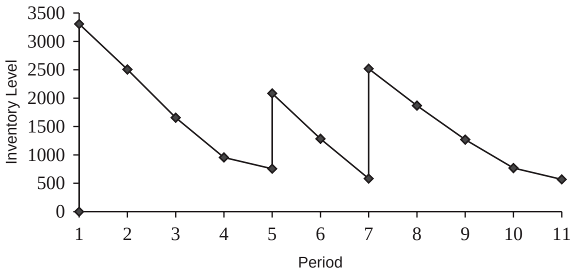

Fig. 1. Replenishment policy of BT.

programming model presented in Section 4.2. The results are presented in Table 4 and Fig. 2. It is seen that this optimal solution has four rather than three replenishments which occur in periods 1, 3, 5 and 8.

Figs. 1 and 2 show the differences between the near optimal BT and the optimal TK policies under the "static–dynamic uncertainty" strategy. The probability of shortage occurring in each time period is given for both policies in Table 5. Since the probability of a shortage occurring never exceeds 5% in any period, it is clear from Table 5 that both policies satisfy the service-level constraints that at the end of each period the probability that the net inventory will not be negative is at least $\alpha = 95\%$. The heuristic BT method costs 1.55% more than the expected cost of the optimal policy.

The opening stocks in Fig. 1 (the BT heuristic policy) for periods 1, 5 and 7 when replenishments arrive are values given by the policy. The opening stocks for intermediate periods are expected values, assuming that the period demands occur at their expected value. It is the same for Fig. 2 (the optimal TK solution) where replenishments arrive in periods 1, 3, 5 and 8. As stated the opening stock at the time of a replenishment is known and set by the policy. However, the sizes of the replenishment orders are determined dynamically as demand is realised. A given period's lot-size, $R_t - I_{t-1}$, cannot be found until the realised demand is known. In other words one waits for the demands to become known, having decided in advance how this knowledge will be used. Similarly the actual opening stocks realised in the intermediate periods would depend on the realised demands. On the assumption of independence of successive period demands, Figs. 1 and 2 show what the average opening stock would be if the actual 10 period situation were repeated many times.

In the above example, the unit variable cost has been set to zero. At this point, the effect of having unit variable cost in the problem on the inventory policy should be investigated. To do so, the above problem is solved again for $v = 4$ using both BT and TK. Since BT heuristic ignores the unit variable cost in both the "static uncertainty" and "static–dynamic uncertainty" models, the replenishment periods, which were {1,5,7}, and the calculated expected lot-sizes remain the same. Hence, Fig. 1 still depicts the replenishment policy for BT. On the other hand, TK approach replaces the replenishment periods $\{1,3,5,8\}_{v=0}$ with $\{1,3,5,7,9\}_{v=4}$. Fig. 3 shows the revised replenishment policy. A comparison of the statistics for both approaches is given in Table 6.

From Table 6, the difference between the expected total inventory costs of BT and TK solutions for $v = 0$ is \$300 (\$19,704–\$19,404). The introduction of unit variable cost, $v = 4$, increases this difference to \$939 (\$45,975–\$45,036).

Table 4
Replenishment policy of TK

Period (t)	1	2	3	4	5	6	7	8	9	10
Order-up-to-level ($E\{R_t\}_{\delta_t=1}$)	2290	—	1299	—	2833	—	—	1742	—	—
Exp. opening inv. ($E\{R_t\}$)	2290	1490	1299	599	2833	2033	1333	1742	1142	642
Exp. closing inv. ($E\{I_t\}$)	1490	640	599	399	2033	1333	683	1142	642	442
Total expected cost	19,404									

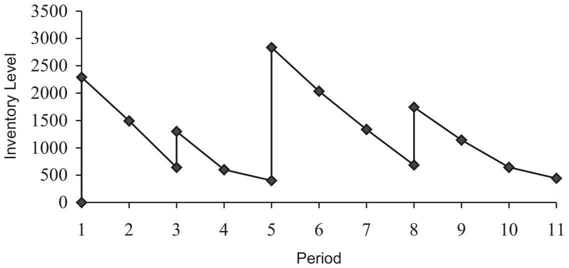

Fig. 2. Replenishment policy of TK.

Table 5
Probability of shortage (%)

Period	1	2	3	4	5	6	7	8	9	10
BT policy	0.0	0.0	1.8	5.0	0.0	5.0	0.0	0.0	1.7	5.0
TK policy	0.0	5.0	0.5	5.0	0.0	0.0	5.0	0.0	0.7	5.0

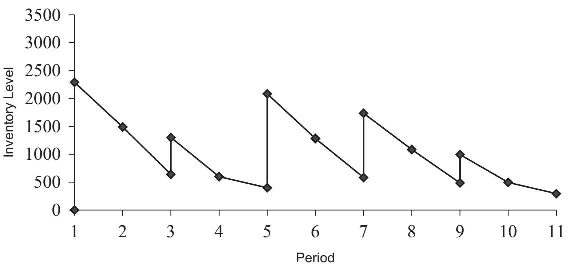

Fig. 3. Replenishment policy of $TK_{v=4}$.

6. Numerical comparisons

In the illustrative example in the previous section, it is seen that the Bookbinder and Tan heuristic costs 1.55% (2.09%) more than the optimal solution for $v = 0$ ($v = 4$). It is useful to gain some idea of how well it would perform over a wider set of examples. This section presents the results of a cost comparison of the BT heuristic with the optimal solution, the TK model, for a wide range of problems. The planning horizon is set to 20 periods with no initial inventory in all the experiments. The service level is set at a constant $\alpha = 0.95$ for all periods and the ordering cost $a = \$1000$ per order. The unit variable cost is ignored. The problems selected are defined in terms of:

- The holding cost, which will affect the average number of periods covered by an order, the values being \$1, \$2, \$3, \$4, \$5, \$7.5 or \$15 per unit per period.
- The coefficient of variation, showing the effect of the size of random variation in demand about the mean. The values selected were 1/3, 1/4, 1/5 and 1/10 and were the same for each period's demand.
- The pattern of the mean demands over time. There are four different demand patterns taken from Berry (1972), as given in Figs. 4a–d. These range from a constant level, through a continuous sinusoidal changes to a very erratic pattern. Demand is normally distributed about the forecast value under the non-stationarity assumption.

The number of test problems, generated for seven different holding costs, four different coefficient of variations and four different mean data sets, therefore totals to 112. For each test problem the percentage cost difference between the BT heuristic and the optimal solution (TK method) is calculated and the results are listed in Table 7.

Table 6
A comparison of $v = 0$ and 4 cases

	$BT_{v=0}$	$BT_{v=4}$	$TK_{v=0}$	$TK_{v=4}$
Replenishment periods	{1,5,7}	{1,5,7}	{1,3,5,8}	{1,3,5,7,9}
E(average inventory level)	1520.4	1520.4	1240.3	1035.4
E(average buffer stock level)	64.5	64.5	54.5	48.0
Total E(order quantity)	6568	6568	6442	6295
E(total cost)	19,704	45,975	19,404	45,036

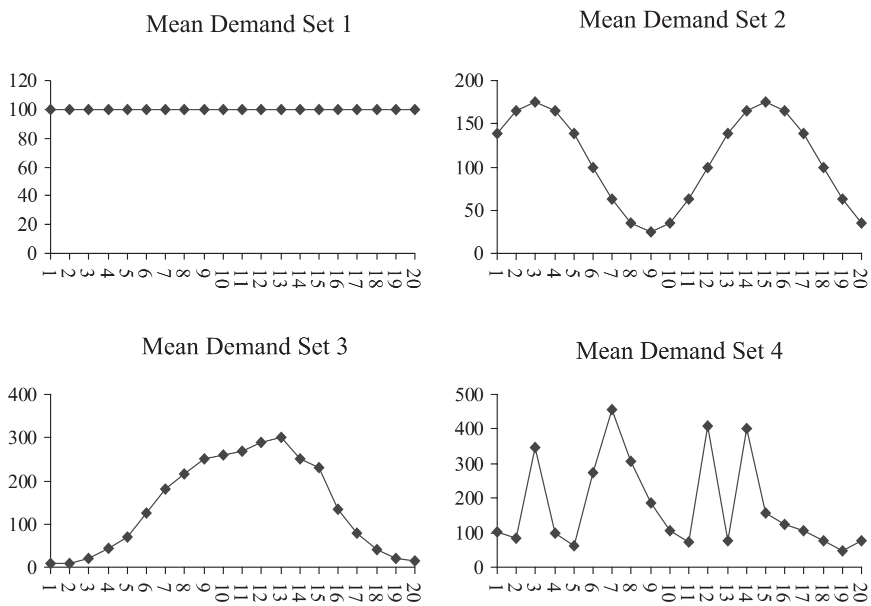

Fig. 4. Demand patterns.

The results indicate that as the demand becomes more erratic, i.e. more non-stationary, either in terms of the mean demand pattern or the size of the coefficient of variation, the cost penalty above the optimal solution from using the BT heuristic tends to become larger. For the stationary case, data set 1, there are 78% of cases where the BT heuristic gives the optimal solution. The average penalty cost above the optimal solution incurred is 0.5%, whilst the worst penalty cost is 3.3%. For data set 2, the sinusoidal pattern to demand over time, the BT heuristic gives the optimal solution for 14% of cases. The average cost penalty is again only 1.3%, whilst the worst penalty is 3.8%, very similar to case 1. For the more erratic almost sinusoidal pattern of data set 3 the BT heuristic gives the optimal solution in 36% of cases. The average cost penalty is 1.24%, similarly to the first two data sets. However, the worst cost penalty is significantly higher at 9.5%. In the highly erratic demand data set 4, the BT heuristic is only optimal in 11% of cases. The average penalty cost incurred is 2.4%, so much higher than the other three cases. The worst penalty is 9.3%, much the same as for data set 3.

The percentage cost difference between BT and the optimal solution tends to become smaller as the coefficient of variation decreases. This is easy to explain. The suboptimality of the BT heuristic arises from the first step of the method. Since the replenishment periods are fixed under the "static uncertainty" strategy and the dynamic nature of

Table 7
Percentage cost increases of BT above TK solutions (for $v = 0$)

	σ/μ	$h=1$	$h=2$	$h=3$	$h=4$	$h=5$	$h=7.5$	$h=15$
Set 1	1/3	0.0	0.0	2.6	0.0	0.0	3.3	0.0
	1/4	0.0	2.5	2.6	0.0	0.0	0.2	0.0
	1/5	0.0	0.0	1.8	0.0	0.0	0.0	0.0
	1/10	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Set 2	1/3	1.9	3.5	3.8	2.2	3.3	1.6	0.1
	1/4	1.0	2.5	2.6	1.7	2.3	0.8	0.0
	1/5	0.5	2.0	2.0	1.3	1.4	0.3	0.0
	1/10	0.0	0.4	0.8	0.4	0.0	0.1	0.0
Set 3	1/3	0.8	3.3	9.5	0.1	1.1	0.7	0.0
	1/4	0.1	2.0	6.8	0.0	0.5	0.5	0.0
	1/5	0.0	1.1	5.1	0.0	0.1	0.4	0.0
	1/10	0.0	0.0	2.3	0.0	0.1	0.1	0.0
Set 4	1/3	5.6	1.0	1.9	5.9	9.3	2.1	0.2
	1/4	4.3	0.3	0.7	3.8	7.5	1.6	1.0
	1/5	3.5	0.0	1.4	2.4	6.1	1.4	0.9
	1/10	1.5	0.0	0.6	0.0	2.6	0.4	0.4

the problem is ignored, in general the replenishment schedule is not optimal. Therefore, a decrease in the coefficient of variation improves the performance of the "static uncertainty" approach and yields a better replenishment schedule. It should be noted that if the replenishment schedule is optimal, then the second step of the method produces the optimal solution for the adjustments to be made in the lot-sizes.

In all four demand data sets it can be seen clearly that the cost penalty increases as the coefficient of variation increases. There is no such simple pattern as the holding cost increases. For example, in the deterministic demand data set 1, the BT heuristic gives the optimal solution for the smallest, middle and largest holding costs but not for all of the others. This is mainly due to the discrete nature of the model used. The replenishment schedule obtained by the "static uncertainty" model will remain unaltered over some range for h. As a result of this, for a specific replenishment schedule the cost performance deteriorates as h approaches to the limits of its range. One observes that at the limits the "static uncertainty" model gives new replenishment policies and the cost performance of the heuristic changes dramatically.

If the unit variable cost is ignored then the stochastic lot-sizing problem can be modelled as a stochastic form of the Wagner–Whitin problem and solved by a shortest route algorithm, where the arc cost (i,j) corresponds to the minimum cost of placing an order in period i to cover the next $j-i+1$ periods and satisfy the service-level constraint. However, the shortest route approach considers the replenishment cycles independent of each other, and as a consequence of this the closing stock of one cycle may be above the opening stock for the next cycle. It is clear that this gives a negative replenishment level, and therefore, is an infeasible solution. Such infeasible solutions are observed particularly in the erratic demand case, demand data set 4. In this case, the shortest route method gives infeasible solution for all problems when σ/μ is in $\{1/3,1/4,1/5\}$ and $h \geqslant 14$. This does not happen with the Bookbinder and Tan heuristic nor the optimal mixed integer programming model presented in this paper. Moreover, if the unit variable cost cannot be ignored, then it is not possible to treat replenishment cycles independently and therefore to apply the shortest route algorithm.

All experiments were done on a 1.2 GHz Pentium III, 512 MB RAM machine. The general-purpose solver CPLEX 8.0 (2002) is used with the default settings and in all cases the optimal solution to the mixed integer programming model is found in less than a quarter of a minute. Although the BT heuristic provides an almost immediate solution, since the computation time required by the TK model is not excessive, it may be worth having the optimal solution.

7. Conclusions

In this paper, the stochastic dynamic lot-sizing problem with service-level constraints has been modelled under the "static–dynamic uncertainty" strategy of Bookbinder and Tan. A mixed integer programming model for the approach has been formulated. This gives the optimal solution allowing the simultaneous determination of the number and timing of the replenishments and the information necessary to determine the size of the replenishment orders, from the replenishment levels for the periods when stock reviews will take place. Unlike the Bookbinder and Tan model this new MIP model includes a unit variable purchasing/production cost. This model allows an estimation of the accuracy of the Bookbinder and Tan heuristic for solving the "static–dynamic uncertainty" approach to be made, by setting the unit cost equal to zero. If the demand data sets, coefficients of variation and relative holding cost to ordering values used in the numerical experiments could be regarded as typical of what occurs in practice we could conclude that the BT heuristic has a cost performance that is close to the optimal solution. Overall, it gave the optimal solution in 36% of cases and had an average penalty cost of 1.34%. Compared to the optimal solution, which is a mixed integer programming model, the solution times are fast. However, the penalty cost for the BT heuristic will be higher for a non-zero unit purchase/production cost. Moreover, in such cases the problem cannot be modelled as a stochastic form of the Wagner–Whitin problem, treating the replenishment cycles independently and applying the shortest route algorithm.

Although, it has been assumed that the replenishment lead-time is zero, it is possible to extend the model for the non-zero replenishment lead-time situation without any loss of generality. A similar model, incorporating the shortage cost, in place of service-level constraints is currently being developed by the authors. Further work could be invested in evaluating the performance of the optimal "static–dynamic uncertainty" strategy in the rolling horizon environment compared to Bookbinder and Tan's heuristic.

References

Askin, R.G., 1981. A procedure for production lot-sizing with probabilistic dynamic demand. AIIE Transactions 13, 132–137.

Berry, W.L., 1972. Lot sizing procedures for requirements planning systems: A framework for analysis. Production and Inventory Management Journal 13, 19–34.

Birge, J.R., Louveaux, F., 1997. Introduction to Stochastic Programming. Springer, New York.

Bookbinder, J.H., Tan, J.Y., 1988. Strategies for the probabilistic lot-sizing problem with service level constraints. Management Science 34, 1096–1108.

Chen, F.Y., Krass, D., 2001. Inventory models with minimal service level constraints. European Journal of Operational Research 134, 120–140.

Davis, T., 1993. Effective supply chain management. Sloan Management Review 34, 35–46.

De Bodt, M.A., Gelders, L.F., Van Wassenhove, L.N., 1984. Lot-sizing under dynamic demand conditions: A review. Engineering Costs and Production Economics 8, 165–187.

Gavirneni, S., Tayur, S., 2001. An efficient procedure for non-stationary inventory control. IIE Transactions 33, 83–89.

Iida, T., 1999. The infinite horizon non-stationary stochastic inventory problem: Near myopic policies and weak ergodicity. European Journal of Operational Research 116, 405–422.

ILOG CPLEX 8.0, 2002. User's Manual, ILOG.

Karni, R., 1981. Integer linear programming formulation of the material requirements planning problem. Journal of Optimization Theory and Applications 35, 217–230.

Kimms, A., 1997. Multi-Level Lot-Sizing and Scheduling Methods for Capacitated Dynamic and Deterministic Models. Physica Verlag Series on Production and Logistics. Springer, Berlin.

Kuik, R., Salomon, M., Van Wassenhove, L.N., 1994. Batching decisions: Structure and models. European Journal of Operational Research 75, 243–263.

Martel, A., Diaby, M., Boctor, F., 1995. Multiple items procurement under stochastic nonstationary demands. European Journal of Operational Research 87, 74–92.

Potts, C.N., Van Wassenhove, L.N., 1992. Integrating scheduling with batching and lot-sizing: A review of algorithms and complexity. Journal of the Operational Research Society 43, 395–406.

Rosling, K., 1999. The square-root algorithm for single-item inventory optimization. Working Paper, Lund University.

Silver, E.A., 1978. Inventory control under a probabilistic time-varying demand pattern. AIIE Transactions 10, 371–379.

Sobel, M.J., Zhang, R.Q., 2001. Inventory policies for systems with stochastic and deterministic demand. Operations Research 49, 157–162.

Sox, C.A., 1997. Dynamic lot-sizing with random demand and non-stationary costs. Operations Research Letters 20, 155–164.

Van Houtum, G.J., Zijm, W.H.M., 2000. On the relation between cost and service models for general inventory systems. Statistica Neerlandica 54, 127–147.

Wagner, H.M., Whitin, T.M., 1958. Dynamic version of the economic lot size model. Management Science 5, 89–96.

Wemmerlov, U., 1989. The behaviour of lot-sizing procedures in the presence of forecast errors. Journal of Operations Management 8, 37–47.

Zangwill, W.I., 1968. Minimum concave cost flows in certain networks. Management Science 14, 429–450.