

A multiperiod stochastic production planning and sourcing problem with service level constraints*

Işıl Yıldırım¹, Barış Tan², and Fikri Karaesmen¹

¹ Department of Industrial Engineering, Koç University, Rumeli Feneri Yolu, Sarıyer, Istanbul, Turkey (e-mail: isil.yildirim@insead.edu.tr;fkaraesmen@ku.edu.tr)

² Graduate School of Business, Koç University, Rumeli Feneri Yolu, Sarıyer, Istanbul, Turkey (e-mail: btan@ku.edu.tr)

Abstract. We study a stochastic multiperiod production planning and sourcing problem of a manufacturer with a number of plants and/or subcontractors. Each source, i.e. each plant and subcontractor, has a different production cost, capacity, and lead time. The manufacturer has to meet the demand for different products according to the service level requirements set by its customers. The demand for each product in each period is random. We present a methodology that a manufacturer can utilize to make its production and sourcing decisions, i.e., to decide how much to produce, when to produce, where to produce, how much inventory to carry, etc. This methodology is based on a mathematical programming approach. The randomness in demand and related probabilistic service level constraints are integrated in a deterministic mathematical program by adding a number of additional linear constraints. Using a rolling horizon approach that solves the deterministic equivalent problem based on the available data at each time period yields an approximate solution to the original dynamic problem. We show that this approach yields the same result as the base stock policy for a single plant with stationary demand. For a system with dual sources, we show that the results obtained from solving the deterministic equivalent model on a rolling horizon gives similar results to a threshold subcontracting policy.

Keywords: Stochastic production planning – Service level constraints – Subcontracting

* The authors are grateful to Yves Dallery for his ideas, comments and suggestions on the earlier versions of this paper.

Correspondence to: F. Karaesmen

1 Introduction and motivation

In this study, we consider a manufacturer that supplies products to a retailer. The manufacturer has a number of production sources that are either its own plants or its subcontractors. Each source has a different production cost, capacity, and lead time. The demand for each product in each period is random. The manufacturer has to meet the demand for multiple products taking into account the service level requirements set by the retailer.

In the production planning and the sourcing problem, the manufacturer's decision variables are how much to produce, when to produce, where to produce, and how much inventory to carry in each period. The objective is to minimize its total production and inventory carrying costs during the planning horizon subject to the service level requirements and other possible constraints.

This problem is motivated by the problems faced by suppliers of lean retailers in the textile-apparel-retail channel (Abernathy et al., 1999). Namely, adoption of lean retailing practices force suppliers of lean retailers to adopt new strategies to respond quickly to changing demand effectively. Using subcontractors emerge as a viable alternative to increase production capacity temporarily when it is needed. Additional cost of subcontracting can be justified by lowering inventories and improving the service. However, deciding on where to produce and how much to produce is a challenging task especially when the demand is volatile. A qualitative discussion of this problem can be found in Abernathy et al. (2000). Figure 1 below depicts the system which motivates this study.

We propose a solution methodology that is based on solving a deterministic mathematical problem at each time period on a rolling horizon basis. Randomness in the problem that comes from uncertain demand and service level constraints are integrated in a deterministic mathematical program by adding a number of additional linear constraints similar to the approach proposed by Bitran and Yanasse (1984). We propose using this approach to address the more relevant but also more

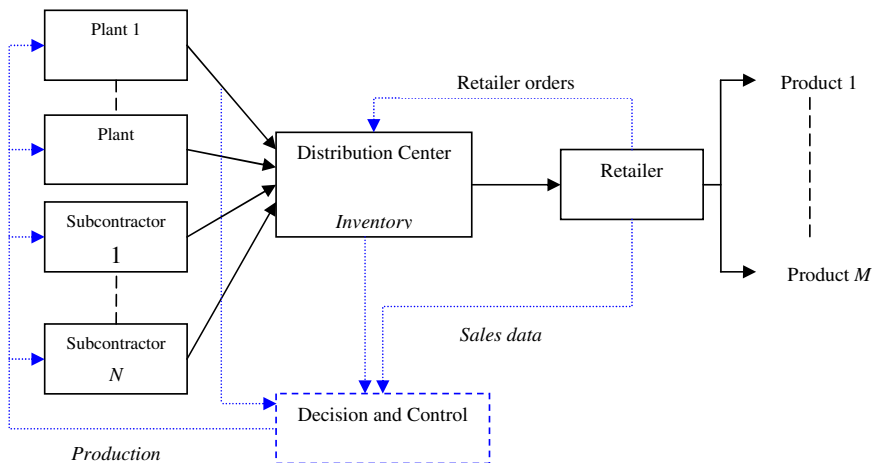


Fig. 1. A manufacturer with multiple plants that sells multiple products to a retailer

difficult dynamic problem where decisions can be updated over time. Since the equivalent deterministic problem is a well-structured mathematical programming problem, the proposed methodology can easily be integrated with the Advanced Planning and Optimization tools, such as the products of i2, Manugistics, etc., that are commonly used in practice.

The organization of the remaining parts of the paper is as follows: In Section 2, we review the literature on mathematical-programming-based stochastic production planning methodologies. The particular stochastic production planning and sourcing problem we investigate is introduced in Section 3. Section 4 presents the proposed solution methodology that is based on solving the deterministic equivalent problem at each time step on a rolling horizon basis. The performance of the rolling horizon approach is evaluated by considering a number of special cases in Section 5. Finally, conclusions are presented in Section 6.

2 Literature review

The classical deterministic production planning problem, its mathematical programming formulations and solution methodologies have received a lot of attention for many years (see Hax and Candea, 1984 for a number of well-known models). In this section, we only review the literature directly related to mathematical programming based approaches for stochastic production planning problems.

Bitran and Yanasse (1984) deal with a similar stochastic production planning problem with a service level requirement. They provide non-sequential (static) and deterministic equivalent formulations of the model and propose error bounds between the exact solution and the proposed approach. Their main focus is on the solution of the static problem, i.e., the solution at time zero for the whole planning horizon.

Bitran, Haas and Matsudo (1986) present a model that is motivated by a case in the consumer electronics and textile and apparel industry. In this model, the stochastic problem is transformed into a deterministic one by replacing the random demand with their average values. Then, the solution of the transformed problem provides answers to the questions of what to produce and when to produce. The complete solution is obtained by determining how much to produce from a newsboy-type formulation based on the solution of the deterministic problem.

Feiring and Sastri (1989) focus on production smoothing plans with rolling horizon strategies and confidence levels for the demand, which are set by the production planners. The probabilistic constraints in the demand-driven scheduling model are revised by Bayesian procedures and are transformed into deterministic constraints by inverse transformations of normally distributed demand.

Zäpfel (1996) claims that MRP II systems can be inadequate for the solution of production planning problems with uncertain demand because of the insufficiently supported aggregation/disaggregation process. The paper then proposes a procedure to generate an aggregate plan and a consistent disaggregate plan for the Master Production Schedule.

Kelle, Clendenen and Dardeau (1994) extend the economic lot scheduling problem for the single-machine, multi-product case with random demands. Their ob-

jective is to find the optimal length of production cycles that minimizes the sum of set-up costs and inventory holding costs per unit of time and satisfies the demand of products at the required service levels.

Clay and Grossman (1997) focus on a two-stage fixed-recourse problem with stochastic Right-Hand-Side terms and stochastic cost coefficients and propose a sensitivity-based successive disaggregation algorithm.

Sox and Muckstadt (1996) present a model for the finite-horizon, discrete-time, capacitated production planning problem with random demand for multiple products. The proposed model includes backorder cost in the objective function rather than enforcing service level constraints. A subgradient optimization algorithm is developed for the solution of the proposed model by using Lagrangian relaxation and some computational results are provided.

Beyer and Ward (2000) report a production and inventory problem of Hewlett-Packard's Network Server Division. The authors propose a method to incorporate the uncertainties in demand in an Advanced Planning System utilized by Hewlett-Packard.

Albritton, Shapiro and Spearman (2000) study a production planning problem with random demand and limited information and propose a simulation based optimization method. Qui and Burch (1997) study a hierarchical production planning and scheduling problem motivated by the fibre industry and propose an optimization model that uses logic of expert systems.

Van Delft and Vial (2003) consider multiperiod supply chain contracts with options. In order to analyze the contracts, they propose a methodology to formulate the deterministic equivalent problem from the base deterministic model and from an event tree representation of the stochastic process and solve the stochastic linear program by discretizing demand under the backlog assumption.

For the textile-apparel-retail problem discussed in Abernathy et al. (2000), a simulation model has also been developed (Yang et al., 1997). Then a simulation-based optimization technique that is referred as ordinal optimization, has been used to determine the parameters of a production and inventory control policy that gives a good-enough solution approximately (Yang et al., 1997; Lee, 1997). However, one needs to set a specific production and inventory control policy in the simulation. In addition to the difficulty of setting a plausible policy in a complicated case, as the number of sources and products increase, the number of parameters to be optimized also increases. As a result, finding an approximate solution requires a considerable time.

Simplified versions of the sourcing problem studied in this paper have been investigated in the past by using stochastic optimal control (Bradley, 2002; Tan and Gershwin, 2004; Tan, 2001). Bradley (2002) considers a system with a producer and a subcontractor and discrete flow of goods. In an M/M/1 setting without the service level requirements, he proves that the optimal control policy structure is a dual-base stock policy. In this policy when the number of customers in the queue reaches a certain level, then new incoming customers are sent to the subcontractor. When there are no customers waiting in the queue, then the producer continues production until a certain threshold is reached.

In Tan (2001) and Tan and Gershwin (2004), a producer with a single subcontractor is formulated with continuous flow of goods without the service level requirements. They also show that a threshold-type policy is optimal to decide when and how to use a subcontractor. In the threshold policy, the subcontractor is used when the inventory or the backlog is below a certain threshold level.

Our paper uses the idea of incorporating randomness in a deterministic mathematical program that is used in many of the above studies in different formats. We utilize the approach proposed by Bitran and Yanasee (1984) that shows the equivalence for the static problem. In contrast to this study where the main objective is determining error bounds for the optimal cost in the non-sequential case, our main focus is generating a production and sourcing plan, i.e. determining the values of the decision variables in the sequential (dynamic) problem where sourcing decisions are made (or updated) dynamically over time. We also compare the approximate solution of the dynamic problem with certain benchmark policies. Since the exact optimal solution of the dynamic problem is not known, we use two different benchmarks. It is proven that for a single source with lead time, the proposed approach yields the same production policy as the optimal base stock policy. For a dual-source, e.g. a producer with a subcontractor, a threshold-type subcontracting policy suggested by Bradley (2002), Tan (2001), Tan and Gershwin (2004) is utilized as a benchmark. After adopting the threshold policy to a more generalized case with lead time and service-level requirements, it is observed that the proposed approach yields very similar results to the threshold-based benchmark in the numerical examples considered.

3 Stochastic multiperiod sourcing problem with service level constraints

Assume that there is a single product and N different production sources (plants and subcontractors). The demand for this specific product at time t , d_t is random. The main decision variables are the production quantities at each production source at time t , $X_{i,t}$, $i = 1, \dots, N$. The inventory level at the end of time period t is denoted by I_t . The number of periods in the planning horizon is T . The inventory holding cost per unit per unit time is h_t and the production cost at production source i at time t is $c_{i,t}$.

Constraints on the performance (related to backorders) of the system are imposed by requiring service levels. The frequently used *Type 1 Service Level* is defined to be the fraction of periods in which there is no stock out. It can be viewed as the plant's no-stock-out frequency. This service level measures whether or not a backorder occurs but is not concerned with the size of the backorder. In this study, we consider a *Modified Type 1 Service Level* requirement. The *Modified Type 1 Service Level* forces the probability of having no stock out to be greater than or equal to a service level requirement in each period. The service level requirement in period t is denoted by α_t .

The Stochastic Production Planning and Sourcing Problem (SP) is defined as:

$$Z^*(SP) = \text{Min}E \left[\sum_{t=1}^T \left(h_t(I_t)^+ + \sum_{i=1}^N c_{i,t}X_{i,t} \right) \right]$$

subject to

$$I_t = I_{t-1} + \sum_{i=1}^N X_{i,t} - d_t, \quad t = 1, \dots, T; \tag{1}$$

$$P \{I_t \geq 0\} \geq \alpha_t, \quad t = 1, \dots, T. \tag{2}$$

$$X_{i,t} \geq 0, \quad i = 1, \dots, N \quad t = 1, \dots, T. \tag{3}$$

where $(I_t)^+ = \text{Max} \{0, I_t\}$, $t = 1, \dots, T$.

The objective of the problem is to minimize the total expected cost, which is the expected value of the sum of the inventory holding and production costs in the planning horizon. The first constraint set defines the inventory balance equations for each time period. The next constraint imposes the service level requirement for each period. Finally, the last constraint states that the production quantities cannot be negative.

This formulation can easily be extended to multiple products and production sources with lead times. Moreover different service level definitions can also be considered by following the same approach.

4 An approximate solution procedure based on a rolling horizon procedure

The solution of the above problem at time 0 for the planning horizon $[0, T]$ is referred as the static solution. The static solution is obtained by using the available information about the distribution of demand in the future periods and the initial inventory. A policy that sets (or updates) the future production quantities $X_{i,t}$ at time t based on the information available at that time, e.g., demand realizations, demand distributions in the future periods, and current inventory levels, is referred to as the dynamic solution.

In theory, the optimal policy which determines production quantities based on actual state information may be obtained by solving the stochastic dynamic program associated with this problem. In practice, however, there are several problems with the stochastic dynamic programming solution. First, the well-known curse of dimensionality makes numerical solutions challenging even for relatively small problems. Second, it is difficult to integrate constraints on the trajectory of the underlying stochastic processes such as service level requirements in inventory models. Therefore, we propose a rolling-horizon approach that is based on solving the static problem at each time period based on the available information. This, however, requires solving the static problem repeatedly which requires a transformation explained below.

4.1 Deterministic equivalent formulation for the static solution

Although obtaining the optimal dynamic solution is, in general, not tractable, the static solution can relatively easily be obtained by using deterministic mathematical programming as suggested by Bitran and Yanasse (1984).

In particular, Bitran and Yanasse show that the (Modified Type 1) service level constraint can be transformed into a deterministic equivalent constraint by specifying certain minimum cumulative production quantities that depend on the service level requirements.

To summarize this approach, let l_t denote the (deterministic equivalent) minimum cumulative production quantity in period t which is calculated by solving the probabilistic inequality:

$$P \left\{ \sum_{\tau=1}^t d_{\tau} \leq l_t \right\} = \alpha_t, \quad t = 1, \dots, T \text{ for } l_t(t = 1, \dots, T)$$

that yields

$$l_t = F_t^{-1}(\alpha_t), \quad t = 1, \dots, T$$

where $F_t(\cdot)$ is the cumulative distribution function of the random sum $\sum_{\tau=1}^t d_{\tau}$. Then the probabilistic constraint $P\{I_t \geq 0\} \geq \alpha_t, t = 1, \dots, T$ can be expressed equivalently by:

$$\sum_{\tau=1}^t \sum_{i=1}^N X_{i,\tau} + I_0 \geq l_t, \quad t = 1, \dots, T \tag{4}$$

Now, the deterministic equivalent problem with service level constraints that has been mentioned in the previous sections can be modeled as below (Bitran and Yanasse, 1984):

Deterministic Equivalent Problem (DEP):

$$Z^*(DEP) = \text{Min} \sum_{t=1}^T \left(h_t(I_0 + \sum_{\tau=1}^t \sum_{i=1}^N X_{i,\tau}) + \sum_{i=1}^N c_{i,t} X_{i,t} \right)$$

subject to

$$\sum_{\tau=1}^t \sum_{i=1}^N X_{i,\tau} + I_0 \geq l_t, \quad t = 1, \dots, T \tag{5}$$

$$X_{i,t} \geq 0, \quad i = 1, \dots, N \quad t = 1, \dots, T. \tag{6}$$

The optimal decision variable values in DEP are the same as the ones in the solution of SP at time 0.

The static solution is obtained by transforming the stochastic problem into a deterministic one and then solving the resulting mathematical program. The rolling horizon approach repeats this procedure by using the available information at each time period until time T .

5 Performance of the rolling horizon solution

It is known that the rolling-horizon approach yields good results for a number of dynamic optimization problems. In some special cases, the rolling horizon method may even yield the optimal solution. In this section, we evaluate the performance of the proposed method by comparing it to certain benchmark policies in two commonly encountered special cases in production planning.

5.1 A single source problem with stationary demand

We start with the special case of a single production source. When there is only one source, the objective function includes only the holding cost (since the expected total production costs must equal the total expected demand over the planning horizon). In this case, we use the base stock policy as the benchmark policy. The base stock policy is widely known and utilized in many applications. In addition, it is known to be optimal in a number of related inventory problems. It, therefore, constitutes a natural benchmark for comparison. The base stock policy has a single parameter which is a reorder level and a base lot size of one unit. It aims to maintain a pre-specified target inventory level. Under this policy, the sequence of events is as follows: the system starts with a pre-specified base stock level in the finished goods inventory. The arrival of the customer demand triggers the consumption of an end-item from the inventory and issuing of a replenishment order to the production facility. Using this policy, an order is placed (or the manufacturing facility operates) if and only if the inventory level drops below the base stock level. The comparison of these two models is performed for two cases with and without a lead time.

5.1.1 Single source without lead time

In this first scenario, there is a single product to be produced by a single production facility. It is assumed that the demand of this specific product stays stationary over the planning horizon. We propose that solving the deterministic equivalent model with modified service level constraints on a rolling horizon basis is equivalent to operating the system under the base stock policy. The next proposition establishes this equivalence:

Proposition 1. When the production facility has no lead time and the demand is stationary, using a base stock policy is equivalent to solving the deterministic equivalent model with service level constraints on a rolling horizon basis (either Modified Type 1 or Modified Type 2) in the following way: assume that the base stock level in the base stock policy equals $I_0(BS) = S_1$ and the initial inventory level in the deterministic equivalent problem equals $I_0(DEP) = l_1$. If $S_1 = l_1$, then the equivalent base stock policy gives the same total expected cost value, yields the same production plan and results in the same service level with the deterministic equivalent model with modified service level constraints solved on a rolling horizon basis.

Since this case is a special case of the next one with lead time, the proof of Proposition 1 is not given here but reported in (Yıldırım, 2004).

Corollary 1. The optimal base stock level is equal to l_1 . Equivalently, the base stock level $S_1 = l_1$ ensures that the resulting production plan satisfies the required service levels.

Proof. If the initial inventory level is set to be $S_1 = l_1$, the resulting production plan is the same with that of the base stock policy which starts with a base stock level of $S_1 = l_1$. Although the base stock policy does not guarantee the assurance of the service levels, since we know that the deterministic equivalent model satisfies the

required service levels and the two policies are equivalent, we can say that the base stock level $S_1 = l_1$ ensures that the resulting production plan satisfies the required service levels. Note that $S_1 = l_1$ must be optimal because decreasing the base stock level from l_1 leads to an infeasible solution and increasing it above l_1 would lead to higher average inventory costs and therefore cannot be optimal. \square

Even though a formal proof is lacking, it is highly likely that the base stock policy (with a stationary base stock level) is optimal for the single-plant single-product problem in an infinite horizon setting. Theorem 1 and Corollary 1 establish that for this problem, the rolling horizon approach yields the same solutions as the optimal base stock policy leading us to conclude that the rolling horizon procedure performs optimally in this case.

5.1.2 Single source with lead time

The deterministic equivalent model with service level constraints (DEP) can be extended to a case in which the production facility has a production lead time. Assume that there is a production lead time of LT periods and the initial scheduled receipts are denoted by $SR_t, t = 1, \dots, LT$. Then, the problem can be modeled in the following way:

Deterministic Equivalent Production Planning Problem including Lead Time (DEPLT):

$$Z^*(\text{DEPLT}) = \text{Min} \sum_{t=1}^{LT} \left(h_t(I_0 + \sum_{\tau=1}^t SR_\tau) \right) + \sum_{t=LT+1}^T \left(h_t(I_0 + \sum_{\tau=1}^{LT} SR_\tau + \sum_{\tau=LT+1}^t \sum_{i=1}^N X_{i,\tau-LT}) \right)$$

subject to

$$\sum_{\tau=LT+1}^t X_{\tau-LT} + \sum_{\tau=1}^{LT} SR_\tau + I_0 \geq l_t, \quad t = (LT + 1), \dots, T; \tag{7}$$

$$X_t \geq 0, \quad t = 1, \dots, T. \tag{8}$$

Our main result is as follows:

Proposition 2. When the production facility has a non-negative lead time LT , the demand is stationary and there are no scheduled receipts initially, using a base stock policy is equivalent to solving the deterministic equivalent model with service level constraints on a rolling horizon basis in the following manner: assume that the base stock level in the base stock policy including lead time equals $I_0(\text{BSLT}) = S_2$ and the initial inventory level in the deterministic equivalent model including lead time equals $I_0(\text{DEPLT}) = l_{LT+1}$. If $S_2 = l_{LT+1}$, then the equivalent base stock policy gives the same total expected cost value, yields the same production plan and results in the same service level with the deterministic equivalent model with service level constraints solved on a rolling horizon basis.

Proof. The proof of Proposition 2 is given in the Appendix. \square

5.2 A dual source problem with stationary demand

Since the optimal solution of our dynamic problem is not known, a plausible benchmark is used to evaluate the performance of the proposed approach. We propose a threshold subcontracting model suggested in a number of studies in the literature (Bradley, 2002; Tan, 2001; Tan and Gershwin, 2004). Although the threshold policy is only shown to be optimal under specific assumptions including zero lead time, stationary demand, no service level requirements, etc., we think that it is a reasonable benchmark policy for our problem.

5.2.1 A threshold subcontracting policy

Now we explain the operation of the threshold policy for our benchmark case. We consider a dual source system with an in-house production facility and a subcontractor. We assume that the in-house facility has a capacity of C but the subcontractor has an infinite capacity. There is a lead time of one period. That is, production quantities scheduled at time t become available at time $t + 1$.

The threshold policy is characterized by two threshold levels S and Z . The in-house production facility operates when the inventory level is below S . That is, it starts producing when the inventory level drops below the target level S and stops producing when the inventory level again reaches S . The subcontractor is used when the inventory level decreases to a threshold level of Z .

When the inventory level is below S , but is still above Z , the in-house facility produces to cover the shortfall with respect to S . If there is not sufficient production capacity to cover the whole shortfall, the in-house facility operates at full capacity and the portion of demand that cannot be satisfied is backlogged for the next period.

Let $X_{1,t}$ and $X_{2,t}$ denote the production amounts of the in-house facility and the subcontractor in period t respectively. Then, the production amounts of each production facility in each time period can be determined for the threshold subcontracting model in the following way:

$$X_{1,t} = \text{Min}\{S - Z, S - I_{t-1}, C\}, \quad t = 1, \dots, T; \quad (9)$$

$$X_{2,t} = \text{Max}\{0, Z - I_{t-1}\}, \quad t = 1, \dots, T. \quad (10)$$

The following figure shows the evolution of $X_{1,t}$, $X_{2,t}$ and I_t under this policy for a Poisson arrival of demand with rate 10 and $S = 15$, $Z = 7$, and $C = 8$.

5.2.2 Comparison of the performance of the threshold policy and the rolling horizon approach

The deterministic equivalent model for this case is solved for a rolling horizon of 10 periods repeatedly throughout a planning horizon of 1000 periods. 5000 sample demand streams are generated and the realized inventory levels are integrated in the model accordingly. The production plans and the realized cost values between periods 451 and 550 are observed. All cost values are calculated on a per period basis.

The optimal values of the threshold values S and Z are determined by using a direct simulation-based numerical search. It is assumed that there are 1000 periods

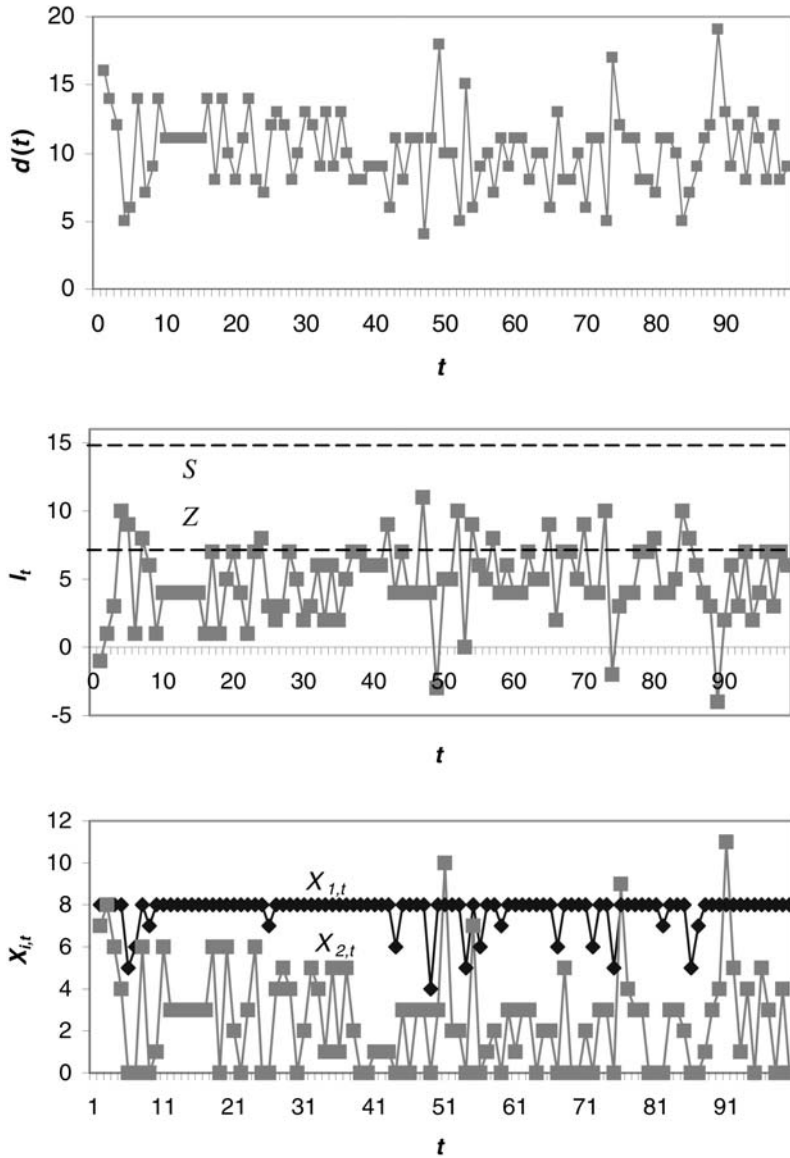


Fig. 2. Sample realization of d_t , $X_{1,t}$, $X_{2,t}$ and I_t under the threshold policy $S = 15$, $Z = 7$, $C = 8$

in the planning horizon and the same 5000 sample demand streams are utilized. The service level requirement is relaxed with the one-sided 95% confidence interval of the simulation result. That is whenever upper confidence level of the observed service level reaches the desired one, this case is accepted as satisfying the service level requirement. The underlying reasoning behind making this modification in service levels is that, the sample size we utilize might not be sufficient enough to make the realized service level equal exactly to the required one. Among the base

Table 1. The possible scenarios for which comparisons are made

Subcontracting cost	Holding cost	In-house production capacity	(Subcontracting cost)/(in-house prod. cost)	(Holding cost)/(in-house prod. cost)	(In-house prod. capacity)/(mean demand)
4	16	8	1	4	0.8
4	16	12	1	4	1.2
4	16	20	1	4	2
6	1	8	1.5	0.25	0.8
6	1	12	1.5	0.25	1.2
6	1	20	1.5	0.25	2
6	4	8	1.5	1	0.8
6	4	12	1.5	1	1.2
6	4	20	1.5	1	2

stock and threshold levels that satisfy the relevant service level requirements, the model aims to find the one with minimum total cost. The calculations are performed for periods between 451 and 550.

For the numerical examples reported below, the order arrivals are governed by a Poisson process with rate 10 products per period. The production cost is assumed to be \$4 per product for the in-house facility. The initial inventory level of the specific product is set to be zero. The service level requirement is set to be 95%.

The comparison between the deterministic equivalent model and the threshold subcontracting model is performed for nine combinations of subcontracting cost to in-house production cost, holding cost to in-house production cost and capacity to mean demand ratios. The combinations of subcontracting costs, holding costs and the in-house production capacities and therefore, the combinations of relevant subcontracting cost to in-house production cost, holding cost to in-house production cost and capacity to mean demand ratios for which the comparisons are made can be observed in Table 1. For each of the problem settings, the base stock and threshold levels observed in the threshold subcontracting model are reported in Table 2.

Note that, in some of the cases, the base stock and threshold pairs are observed to be the same. The reasoning behind this is, these pairs lead to the same average inventory levels and minimum cost values in these settings.

While comparing the two models, total expected cost, average production cost, average inventory holding cost values and the assignment of production to the plants (in percentages) are the key elements we focus on. Table 3 summarizes the total expected cost values of the deterministic equivalent model (DEM) and the threshold subcontracting model (TSM) for the nine different scenarios for each modified service level type.

The below tables display that the deterministic equivalent model gives very close solutions when compared with the threshold subcontracting model for both types of the modified levels. The deterministic equivalent model results in total expected cost values equal to or a little bit larger than those of the threshold subcontracting

Table 2. Base stock and threshold levels observed in each scenario

Subcontracting cost	Holding cost	In-house production capacity	Critical levels	
			Base stock	Threshold
4	16	8	15	7
4	16	12	15	3
4	16	20	15	$-\infty$
6	1	8	17	7
6	1	12	16	0
6	1	20	15	$-\infty$
6	4	8	15	7
6	4	12	15	3
6	4	20	15	$-\infty$

Table 3. The comparison of total expected cost values observed in each scenario

Subcontracting cost	Holding cost	In-house production capacity	Total expected cost		
			DEM	TSM	Percentage difference
4	16	8	121.66	121.66	0.00
4	16	12	121.66	121.66	0.00
4	16	20	121.66	121.62	0.03
6	1	8	49.97	49.89	0.16
6	1	12	46.16	45.65	1.12
6	1	20	45.10	45.10	0.02
6	4	8	65.33	65.33	0.00
6	4	12	61.47	61.47	0.00
6	4	20	60.42	60.40	0.03

model. For our set of numerical experiments, the deterministic equivalent model gives close results to the threshold subcontracting model when the service level requirement is of Modified Type 1.

Tables 4 and 5 display the comparison of average production and holding cost values. As can be seen, the deterministic equivalent model gives similar results to the threshold subcontracting model.

Table 6 summarizes the percentage of production assigned to the in-house production facility for both the deterministic equivalent model and the threshold subcontracting model. The results suggest that the production assignments of the deterministic model follow a similar pattern with the benchmark chosen.

Based on these figures, we can conclude that the proposed deterministic equivalent model solved on a rolling horizon basis performs as well as the threshold subcontracting model solved on a simulation-based optimization technique for the

Table 4. The comparison of average production cost values observed in each scenario

Subcontracting cost	Holding cost	In-house production capacity	Average production cost		
			DEM	TSM	Percentage difference
4	16	8	39.99	39.99	0.00
4	16	12	39.99	39.99	0.00
4	16	20	39.99	39.99	0.00
6	1	8	44.06	44.36	-0.68
6	1	12	41.05	40.24	2.03
6	1	20	40.00	39.97	0.06
6	4	8	44.91	44.91	0.00
6	4	12	41.05	41.05	0.00
6	4	20	40.00	39.99	0.01

Table 5. The comparison of average holding cost values observed in each scenario

Subcontracting cost	Holding cost	In-house production capacity	Average holding cost		
			DEM	TSM	Percentage difference
4	16	8	81.67	81.67	0.00
4	16	12	81.67	81.67	0.00
4	16	20	81.67	81.63	0.05
6	1	8	5.92	5.53	6.91
6	1	12	5.10	5.41	-5.67
6	1	20	5.10	5.10	0.05
6	4	8	20.42	20.42	0.00
6	4	12	20.42	20.42	0.00
6	4	20	20.42	20.41	0.05

Modified Type 1 service level. The total expected cost values of deterministic equivalent models for all nine different cases are equal to or a little bit larger than those of the threshold subcontracting model. However, we cannot reach the same conclusion for the average production and holding cost values. The deterministic equivalent model performs either worse for some cases or better for some other cases when the comparison is based on average production or holding cost values. However, the sum of these two terms, the total expected cost, is equal to a little bit larger than that of the threshold subcontracting model. Moreover, the proportion of production assigned to the in-house facility in the deterministic equivalent model resembles that in the simulation based threshold subcontracting model.

It is worth mentioning that the sample size utilized in the above numerical comparisons, 5000, might not be large enough to satisfy the service level requirements in each time period that the modified service level definitions necessitate.

Table 6. The percentage of production assignments to the in-house production facility observed in each scenario

Subcontracting cost	Holding cost	In-house production capacity	% In-house production	
			Base stock	Threshold
4	16	8	75.45	75.40
4	16	12	94.73	94.70
4	16	20	99.97	100.00
6	1	8	79.76	78.17
6	1	12	94.73	98.78
6	1	20	99.97	100.00
6	4	8	75.45	75.40
6	4	12	94.73	94.70
6	4	20	99.97	100.00

The coefficient of variation in the realized service level values might be larger than expected. To handle this problematic issue, we introduced one-sided confidence intervals. Although the threshold subcontracting model constitutes a lower bound in terms of total expected cost values for our set of numerical examples, it can not be generalized from our examples that the deterministic equivalent model always gives solutions worse than those of the threshold subcontracting model. Nevertheless, the proposed approach seems to give extremely promising results in this particular case as well.

6 Conclusions

In many practical situations, mathematical models of production planning/outsourcing problems have to deal with the randomness in demand. We present a systematic approach that enables the randomness in demand and the desired service levels to be incorporated in a mathematical programming framework.

We show that solving the deterministic equivalent problem on a rolling-horizon basis gives similar results to the performance of the benchmarks. Although the threshold-type policies are conceptually quite intuitive, it is very challenging to determine the optimal threshold levels by using simulation. The proposed algorithm is easier to implement and optimize by using available solvers.

This study can be extended in a number of ways. The same approach can be used to derive results for different service level definitions. Yıldırım (2004) reports preliminary results for Type 2 and Modified Type 2 service levels. The formulation of the multi-product case is also straightforward.

The effects of demand variability, production cost, and the lead time on the production and sourcing plans need further investigation. Since the optimal solution to the general problem is not known for the dynamic case, investigation of the static

case or a stylized model can yield insights regarding the interaction of demand variability, cost, and the lead time.

Appendix

Proof of Proposition 2

We use induction to show that

- i. If the inventory levels at the beginning of the first period are equal, $I_0(\text{BSLT}) = I_0(\text{DEPLT}) = l_{LT+1}$, then production quantities in the first period and the inventory at the end of first period for both policies become equal, i.e. $X_1(\text{BSLT}) = X_1(\text{DEPLT}) = 0$ and $I_1(\text{BSLT}) = I_1(\text{DEPLT}) = l_{LT+1} - d_1$;
- ii. If the inventory levels at the end of period t_1 such that $t_1 \leq LT$ are equal, $I_{t_1}(\text{BSLT}) = I_{t_1}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=1}^{t_1} d_\tau$, then the production quantities in period $(t_1 + 1)$ and the inventory levels at the end of period $(t_1 + 1)$ for both policies become equal; i.e. $X_{t_1+1}(\text{BSLT}) = X_{t_1+1}(\text{DEPLT}) = d_{t_1}$ and $I_{t_1+1}(\text{BSLT}) = I_{t_1+1}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=1}^{t_1+1} d_\tau$.
and
- iii. If the inventory levels at the end of period $(LT+1)$ are equal, $I_{LT+1}(\text{BSLT})=I_{LT+1}(\text{DEPLT})=l_{LT+1} - \sum_{\tau=1}^{LT+1} d_\tau$, then production quantities in period $(LT+2)$ and the inventory levels at the end of period $(LT+2)$ for both policies become equal, i.e. $X_{LT+2}(\text{BSLT}) = X_{LT+2}(\text{DEPLT}) = d_{LT+1}$ and $I_{LT+2}(\text{BSLT}) = I_{LT+2}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=2}^{LT+2} d_\tau$;
- iv. If the inventory levels at the end of period t_2 such that $t_2 \geq LT$ are equal, $I_{t_2}(\text{BSLT}) = I_{t_2}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=t_2-LT}^{t_2} d_\tau$, then the production quantities in period $(t_2 + 1)$ and the inventory levels at the end of period $(t_2 + 1)$ for both policies become equal; i.e. $X_{t_2+1}(\text{BSLT}) = X_{t_2+1}(\text{DEPLT}) = d_{t_2}$ and $I_{t_2+1}(\text{BSLT}) = I_{t_2+1}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=t_2+1-LT}^{t_2+1} d_\tau$.

Assume that the initial inventory levels are equal such that $I_0(\text{BSLT}) = S_2, I_0(\text{DEPLT}) = l_{LT+1}$ and $S_2 = l_{LT+1}$. In the base stock policy, each demand observed is produced in the next period; therefore there is no production in the first period, $X_1(\text{BSLT}) = 0$. In the deterministic equivalent approach, the production quantity in the first period is determined according to the constraint $X_1(\text{DEPLT}) + \sum_{\tau=1}^{LT} SR_\tau(\text{DEPLT}) + I_0(\text{DEPLT}) = X_1(\text{DEPLT}) + 0 + l_{LT+1} \geq l_{LT+1}$ and therefore, $X_1(\text{DEPLT}) \geq 0$. Since the problem is of minimization type, the production quantity in the first period equals zero, i.e. $X_1(\text{DEPLT}) = 0$. Next, a customer demand of d_1 arrives. The end of period inventory for the base stock policy becomes $I_1(\text{BSLT}) = I_0(\text{BSLT}) + SR_1(\text{BSLT}) - d_1 = S_2 + 0 - d_1 = S_2 - d_1$ and the end of period inventory for the deterministic equivalent approach becomes $I_1(\text{DEPLT}) = I_0(\text{DEPLT}) + SR_1(\text{DEPLT}) - d_1 = l_{LT+1} + 0 - d_1 = l_{LT+1} - d_1$. Since we know that $S_2 = l_{LT+1}, I_1(\text{BSLT}) = I_1(\text{DEPLT})$.

In the second period, the base stock policy produces the demand of the first period, i.e. $X_2(\text{BSLT}) = d_1$. At the beginning of the second period, the deterministic equivalent model is rerun since it is solved on a rolling horizon basis.

The demand is assumed to be stationary over the planning horizon. Although solving the model on a rolling horizon basis throughout the planning horizon requires integration of the minimum cumulative production quantities for the number of periods in the rolling horizon into the model, only the minimum cumulative production quantity of period $(LT + 1)$, l_{LT+1} , is fully utilized. The production quantity of the deterministic equivalent model in the second period is determined by $X_2(\text{DEPLT}) + \sum_{\tau=2}^{LT+1} SR_{\tau}(\text{DEPLT}) + I_1(\text{DEPLT}) = X_2(\text{DEPLT}) + X_1(\text{DEPLT}) + I_1(\text{DEPLT}) = X_2(\text{DEPLT}) + 0 + l_{LT+1} - d_1 \geq l_{LT+1}$; therefore, $X_2(\text{DEPLT}) \geq d_1$. In order to minimize the production costs, the production quantity in the second period equals the demand of the first period, i.e. $X_2(\text{DEPLT}) = d_1$. After the arrival of a customer demand of d_2 , the end of period inventory for the base stock policy becomes $I_2(\text{BSLT}) = I_1(\text{BSLT}) + SR_2(\text{BSLT}) - d_2 = S_2 - d_1 - d_2$ and the end of period inventory for the deterministic equivalent approach becomes $I_2(\text{DEPLT}) = I_1(\text{DEPLT}) + SR_2(\text{DEPLT}) - d_2 = l_{LT+1} - d_1 - d_2$. Since $S_2 = l_{LT+1}$, we can say that $I_2(\text{BSLT}) = I_2(\text{DEPLT})$.

Since demand during lead time cannot be satisfied no sooner than $(LT + 1)$ periods of time, the inventory levels at the end of any period t_1 such that $t_1 \leq (LT - 1)$ can be written as $I_{t_1}(\text{BSLT}) = S_2 - \sum_{\tau=1}^{t_1} d_{\tau}$, $I_{t_1}(\text{DEP}) = l_{LT+1} - \sum_{\tau=1}^{t_1} d_{\tau}$ and $S_2 = l_{LT+1}$. In period $(t_1 + 1)$, the base stock policy produces $X_{t_1+1}(\text{BSLT}) = d_{t_1}$. In the deterministic equivalent approach, the production quantity is determined by the constraint $X_{t_1+1}(\text{DEPLT}) + \sum_{\tau=t_1+1}^{t_1+LT} SR_{\tau}(\text{DEPLT}) + I_{t_1}(\text{DEPLT}) = X_{t_1+1}(\text{DEPLT}) + \sum_{\tau=1}^{t_1} X_{\tau}(\text{DEPLT}) + I_{t_1}(\text{DEPLT}) = X_{t_1+1}(\text{DEPLT}) + \sum_{\tau=1}^{t_1-1} d_{\tau} + l_{LT+1} - \sum_{\tau=1}^{t_1} d_{\tau} \geq l_{LT+1}$; therefore, $X_{t_1+1}(\text{DEPLT}) \geq d_{t_1}$. Since the problem is of minimization type, $X_{t_1+1}(\text{DEPLT}) = d_{t_1}$. Then, a customer demand of d_{t_1+1} is observed. The end of period inventory for the base stock policy becomes $I_{t_1+1}(\text{BSLT}) = I_{t_1}(\text{BSLT}) + SR_{t_1+1}(\text{BSLT}) - d_{t_1+1} = S_2 - \sum_{\tau=1}^{t_1} d_{\tau} - d_{t_1+1} = S_2 - \sum_{\tau=1}^{t_1+1} d_{\tau}$ and the end of period inventory for the deterministic equivalent approach becomes $I_{t_1+1}(\text{DEPLT}) = I_{t_1}(\text{DEPLT}) + SR_{t_1+1}(\text{DEPLT}) - d_{t_1+1} = l_{LT+1} - \sum_{\tau=1}^{t_1} d_{\tau} - d_{t_1+1} = l_{LT+1} - \sum_{\tau=1}^{t_1+1} d_{\tau}$. Since $S_2 = l_{LT+1}$, $I_{t_1+1}(\text{BSLT}) = I_{t_1+1}(\text{DEPLT})$.

Similarly, d_{LT+1} is produced by the base stock policy in period $(LT + 1)$, i.e. $X_{LT+1} = d_{LT+1}$. The constraint $X_{LT+1}(\text{DEPLT}) + \sum_{\tau=LT+1}^{2LT} SR_{\tau}(\text{DEPLT}) + I_{LT}(\text{DEPLT}) = X_{LT+1}(\text{DEPLT}) + \sum_{\tau=1}^{LT} X_{\tau} + I_{LT}(\text{DEPLT}) = X_{LT+1}(\text{DEPLT}) + \sum_{\tau=1}^{LT-1} d_{\tau} + l_{LT+1} - \sum_{\tau=1}^{LT} d_{\tau} \geq l_{LT+1}$; i.e. $X_{LT+1}(\text{DEPLT}) \geq d_{LT}$ determines the production quantity of the deterministic equivalent model in period $(LT + 1)$. Then, $X_{LT+1}(\text{DEPLT}) = d_{LT}$. Next, a customer demand of d_{LT+1} arrives. The end of period inventory for the base stock policy becomes $I_{LT+1}(\text{BSLT}) = I_{LT}(\text{BSLT}) + SR_{LT+1}(\text{BSLT}) - d_{LT+1} = S_2 - \sum_{\tau=1}^{LT} d_{\tau} + X_1(\text{BSLT}) - d_{LT+1} = S_2 - \sum_{\tau=1}^{LT} d_{\tau} + 0 - d_{LT+1} = S_2 - \sum_{\tau=1}^{LT+1} d_{\tau}$ and the end of period inventory for the deterministic equivalent approach becomes $I_{LT+1}(\text{DEPLT}) = I_{LT}(\text{DEPLT}) + SR_{LT+1}(\text{DEPLT}) - d_{LT+1} = l_{LT+1} - \sum_{\tau=1}^{LT} d_{\tau} + X_1(\text{DEPLT}) - d_{LT+1} = l_{LT+1} - \sum_{\tau=1}^{LT} d_{\tau} + 0 - d_{LT+1} = l_{LT+1} - \sum_{\tau=1}^{LT+1} d_{\tau}$. Since $S_2 = l_{LT+1}$, $I_{LT+1}(\text{BSLT}) = I_{LT+1}(\text{DEPLT})$.

In period $(LT + 2)$, the base stock policy produces $X_{LT+1}(\text{BSLT}) = d_{LT+2}$. For the deterministic equivalent approach, we know that $X_{LT+2}(\text{DEPLT}) + \sum_{\tau=LT+2}^{2LT+1} SR_{\tau}(\text{DEPLT})I_{LT+1}(\text{DEPLT}) = X_{LT+2}(\text{DEPLT}) + \sum_{\tau=2}^{LT+1} X_{\tau} + I_{LT+1}(\text{DEPLT}) = X_{LT+2}(\text{DEPLT}) + \sum_{\tau=1}^{LT} d_{\tau} + l_{LT+1} - \sum_{\tau=1}^{LT+1} d_{\tau} \geq l_{LT+1}$; i.e. $X_{LT+2}(\text{DEPLT}) \geq d_{LT+1}$ and then, $X_{LT+2}(\text{DEPLT}) = d_{LT+1}$. After the arrival of d_{LT+2} , the following end of period inventory levels are observed $I_{LT+2}(\text{BSLT}) = I_{LT+1}(\text{BSLT}) + SR_{LT+2}(\text{BSLT}) - d_{LT+2} = S_2 - \sum_{\tau=1}^{LT+1} d_{\tau} + X_2(\text{BSLT}) - d_{LT+2} = S_2 - \sum_{\tau=1}^{LT+1} d_{\tau} + d_1 - d_{LT+2} = S_2 - \sum_{\tau=2}^{LT+2} d_{\tau}$ and $I_{LT+2}(\text{DEPLT}) = I_{LT+1}(\text{DEPLT}) + SR_{LT+2}(\text{DEPLT}) - d_{LT+2} = l_{LT+1} - \sum_{\tau=1}^{LT+1} d_{\tau} + X_2(\text{DEPLT}) - d_{LT+2} = l_{LT+1} - \sum_{\tau=1}^{LT+1} d_{\tau} + d_1 - d_{LT+2} = l_{LT+1} - \sum_{\tau=2}^{LT+2} d_{\tau}$. Since we know that $S_2 = l_{LT+1}$, $I_{LT+2}(\text{BSLT}) = I_{LT+2}(\text{DEPLT})$.

Now assume that at the end of any period t_2 such that $t_2 \geq (LT + 1)$, $I_{t_2}(\text{BSLT}) = S_2 - \sum_{\tau=t_2-LT}^{t_2} d_{\tau}$, $I_{t_2}(\text{DEPLT}) = l_{LT+1} - \sum_{\tau=t_2-LT}^{t_2} d_{\tau}$ and $S_2 = l_{LT+1}$. In period $(t_2 + 1)$, $X_{t_2+1}(\text{BSLT}) = d_{t_2}$ and $X_{t_2+1}(\text{DEPLT})$ is determined by the constraint $X_{t_2+1}(\text{DEPLT}) + \sum_{\tau=t_2+1}^{t_2+LT} SR_{\tau}(\text{DEPLT}) + I_{t_2}(\text{DEPLT}) = X_{t_2+1}(\text{DEPLT}) + \sum_{\tau=1}^{t_2} X_{\tau} + I_{t_2}(\text{DEPLT}) = X_{t_2+1}(\text{DEPLT}) + \sum_{\tau=1}^{t_2-1} d_{\tau} + l_{LT+1} - \sum_{\tau=1}^{t_2} d_{\tau} \geq l_{LT+1}$; $X_{t_2+1}(\text{DEPLT}) \geq d_{t_2}$ and since the model is of minimization type $X_{t_2+1}(\text{DEPLT}) = d_{t_2}$. Next, a customer demand of d_{t_2+1} arrives. The end of period inventory levels for both policies become $I_{t_2+1}(\text{BSLT}) = I_{t_2}(\text{BSLT}) + SR_{t_2+1}(\text{BSLT}) - d_{t_2+1} = S_2 - \sum_{\tau=t_2-LT}^{t_2} d_{\tau} + X_{t_2+1-LT}(\text{BSLT}) - d_{t_2+1} = S_2 - \sum_{\tau=t_2-LT}^{t_2} d_{\tau} + d_{t_2-LT} - d_{t_2+1} = S_2 - \sum_{\tau=t_2+1-LT}^{t_2+1} d_{\tau}$ and $I_{t_2+1}(\text{DEPLT}) = I_{t_2}(\text{DEPLT}) + SR_{t_2+1}(\text{DEPLT}) - d_{t_2+1} = S_2 - \sum_{\tau=t_2-LT}^{t_2} d_{\tau} + X_{t_2+1-LT}(\text{DEPLT}) - d_{t_2+1} = S_2 - \sum_{\tau=t_2-LT}^{t_2} d_{\tau} + d_{t_2-LT} - d_{t_2+1} = S_2 - \sum_{\tau=t_2+1-LT}^{t_2+1} d_{\tau}$. Since we know that $S_2 = l_{LT+1}$, $I_{t_2+1}(\text{BSLT}) = I_{t_2+1}(\text{DEPLT})$. This proves our proposition.

References

- Abernathy FH, Dunlop JT, Hammond JH, Weil D (1999) A stitch in time. Oxford University Press, New York
- Abernathy FH, Dunlop JT, Hammond JH, Weil D (2000) Control your inventory in a world of lean retailing. Harvard Business Review Nov-Dec: 169–176
- Albritton M, Shapiro A, Spearman M (2000) Finite capacity production planning with random demand and limited information. Stochastic Programming E-Print Series
- Beyer RD, Ward J (2000) Network server supply chain at HP: a case study. HP Labs Tech Report 2000-84
- Bitran GR, Yanasse HH (1984) Deterministic approximations to stochastic production problems. Operations Research 32: 999–1018
- Bitran GR, Haas EA, Maatsudo (1986) Production planning of style goods with high setup costs and forecast revisions. Operations Research 34(2): 226–236
- Bradley JR (2002) Optimal control of a dual service rate M/M/1 production-inventory model. European Journal of Operational Research (2002) (forthcoming)
- Candea D, Hax AC (1984) Production and inventory management. Prentice-Hall, New Jersey
- Clay RL, Grossman IE (1997) A disaggregation algorithm for the optimization of stochastic planning models. Computers and Chemical Engineering 21(7): 751–774

- Feiring BR, Sastri T (1989) A demand-driven method for scheduling optimal smooth production levels. *Annals of Operations Research* 17: 199–216
- Kelle P, Clendenen G, Dardeau P (1994) Economic lot scheduling heuristic for random demands. *International Journal of Production Economics* 35: 337–342
- Lee LH (1997) Ordinal optimization and its application in apparel manufacturing systems. Ph.D. Thesis, Harvard University, Cambridge, MA
- Qiu MM, Burch EE (1997) Hierarchical production planning and scheduling in a multi-product, multi-machine environment. *International Journal of Production Research* 35(11): 3023–3042
- Sox CR, Muckstadt JA (1996) Multi-item, multi-period production planning with uncertain demand. *IIE Transactions* 28: 891–900
- Tan B, Gershwin SB (2004) Production and subcontracting strategies for manufacturers with limited capacity and volatile demand. *Annals of Operations Research (Special volume on Stochastic Models of Production/Inventory Systems)* 125: 205–232
- Tan B (2002) Managing manufacturing risks by using capacity options. *Journal of the Operational Research Society* 53(2): 232–242
- Van Delft C, Vial J-PH (2003) A practical implementation of stochastic programming: an application to the evaluation of option contracts in supply chains. *Automatica* (to appear)
- Yang MS, Lee LH, Ho YC (1997) On stochastic optimization and its applications to manufacturing. *Lectures in Applied Mathematics* 33: 317–331
- Yıldırım I (2004) Stochastic production planning and sourcing problems with service level constraints. M.S. Thesis, Koç University, Industrial Engineering, Istanbul, Turkey
- Zäpfel G (1996) Production planning in the case of uncertain individual demand extension for an MRP II concept. *International Journal of Production Economics* 46–47: 153–164