Production Planning Problems with Joint Service-Level Guarantee: A Computational Study

Yuchen Jiang*, Juan Xu†, Siqian Shen*, Cong Shi*

* Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48109

{ycjiang, siqian, shicong}@umich.edu

† Industrial and Enterprise Systems Engineering, University of Illinois at Urbana-Champaign, Champaign, IL 61801

juanxu2@illinois.edu

Abstract

We consider a class of single-stage multi-period production planning problems under demand uncertainty. The main feature of our paper is to incorporate a joint service-level constraint to restrict the joint probability of having backorders in any period. This is motivated by manufacturing and retailing applications, in which firms need to decide the production quantities ex ante, and also have stringent service-level agreements. The inflexibility of dynamically altering the pre-determined production schedule may be due to contractual agreement with external suppliers or other economic factors such as enormously large fixed costs and long lead time. We focus on two stochastic variants of this problem, with or without pricing decisions, both subject to a joint service-level guarantee. The demand distribution could be nonstationary and correlated across different periods. Using the sample average approximation (SAA) approach for solving chance-constrained programs, we re-formulate the two variants as mixed-integer linear programs (MILPs). Via computations of diverse instances, we demonstrate the effectiveness of the SAA approach, analyze the solution feasibility and objective bounds, and conduct sensitivity analysis for the two MILPs. The approaches can be generalized to a wide variety of production planning problems, and the resulting MILPs can be efficiently computed by commercial solvers.

Key words: production planning; stochastic programming; mixed integer linear programming; joint service-level constraint; sample average approximation.

History: Received September 2015; revisions received March 2016; accepted May 2016.

1 Introduction

In this paper, we study a class of production planning problems subject to a joint service-level constraint. The problems fall into the category of single-stage multi-period stochastic optimization problems with no recourse decisions (i.e., firms need to plan their production and/or pricing decisions ex ante and cannot change them in subsequent periods). This class of problems is primarily motivated by manufacturing or retailing applications, in which firms have stringent service-level requirements, but do not have sufficient flexibility of altering their decisions due to the related issues, such as contractual agreement with outside suppliers, enormously large fixed costs and long lead time.

We describe several motivating examples for conducting the research. In the first example, CEMEX, a multinational building materials company that often signs contracts with large event organizers. The firm was the key supplier of cement for multiple large events in 2014, including notably a contract with Fédération Internationale de Football Association (FIFA) to supply 28000 tons of cement for the new soccer stadium in Manaus [4]. The firm has to plan production quantities and prices for the committed projects long before starting the project. Moreover, guaranteeing an adequate service-level is absolutely essential for successful completion of these projects. The second example arises in the U.S. automotive industry. The lead time for building a new car model is typically 52 months [8] due to lengthy design and testing cycles, and the lead time for manufacturing an existing car model is typically 10 to 18 months [2]. The fixed cost related to altering a production schedule is also quite high. Due to these reasons, car manufacturers usually make their production plans ahead of the selling season and their planning decisions cannot be easily changed afterwards. Meanwhile, an adequate service-level is important to maintain firms' revenue and goodwill.

To capture the aforementioned applications stylistically, we propose two single-stage multi-period models (with or without pricing decisions), subject to a joint service-level guarantee. The first model concerns the production decisions while the second model concerns both the production and pricing decisions. For the first model, the demands are random, which can be non-stationary and correlated across different periods. Our goal is to minimize the expected total cost, including (linear) production costs, inventory holding costs and backorder penalty costs, subject to a joint service-level constraint over the whole periods to restrict the probability of having unmet demands during the planning horizon. For the second model, we assume a classical additive demand model in which the demand depends linearly on the price plus a random disturbance term (see, e.g., Chen and Simchi-Levi [6]). We consider discrete pricing and continuous pricing options, in which prices are chosen from a given finite set of values or from a bounded price range, respectively. Our goal is to maximize the total expected profit, also subject to a joint service-level constraint.

We remark again that both problems considered in this paper belong to the category of single-stage stochastic optimization problems with no recourse decisions, which should be distinguished from dynamic inventory control problems considered in the literature (see, e.g., Zipkin [26]). The main feature of our models is to incorporate a joint service-level constraint, which is practically relevant but computationally intractable. Our approach employs the *Sample Average Approximation* (SAA) method (see Luedtke and Ahmed [16]) to reformulate our chance constrained problems as mixed-integer linear programming (MILP) models and to compute upper and lower bounds of the optimal objective values as well as feasible solutions with certain confidence levels.

1.1 Relevant Literature

The traditional study of production planning problem has been focused on deterministic models with known demand. Zangwill [25] developed a deterministic lot-sizing model that allowed for backlogged demand and proposed a network approach. They further proposed dynamic programming algorithms to compute optimal planning policies based on network formulations. Pochet and Wolsey [23] studied several strong MILP reformulations of the uncapacitated lot-sizing problem with backlogging. They also described a family of strong valid inequalities that can be effectively used in a cut generation algorithm. Florian et al. [9] studied capacitated lot-sizing problem and showed that the deterministic problem is NP-hard. Recently, Absi et al. [1] studied the single item uncapacitated lot-sizing problem with production time windows, lost sales, early productions and backlogs. They presented MILP formulations of these models and developed dynamic programming algorithms to solve them. González-Ramírez et al. [10] proposed a heuristic algorithm to solve a multi-product, multi-period capacitated lot-sizing problem with pricing, where the deterministic demand was assumed to be linear in price.

For the production planning problem with stochastic demand, Mula et al. [21] reviewed some existing literature of production planning under demand uncertainty. Gupta and Maranas [11] proposed a stochastic programming based approach to incorporate demand uncertainty in midterm production planning. Kazaz [14] studied production planning with random yield and demand using a two-stage stochastic programming approach. Clark and Scarf [7] considered a periodic-review multi-period production planning problem with uncertain demand and they showed the structure of optimal policies via dynamic programming approach.

In this paper, we focus on stochastic variants of production planning problems, subject to a joint service-level constraint. There has been limited literature on this topic, among which Bookbinder and Tan [3] studied a multi-period lot-sizing problem that imposed individual service-level constraint in each period and their demand distributions were known. In contrast, our model considers a joint service-level constraint that poses more computational challenges, and empirical demand samples are given instead of an explicit demand distribution function. We reformulate our problem as an MILP model using the SAA approach, which is based on Monte Carlo simulation of random samples, to approximate the expected value function by the corresponding sample average. Kleywegt et al. [15] studied the convergence rates, stopping rules and computational complexity of the SAA method. They also presented a numerical example for solving the stochastic knapsack problem using the SAA method. Verweij et al. [24] formulated stochastic routing problems using the SAA approach. They applied decomposition and branch-and-cut techniques to numerically solve the approximating problems. Pagnoncelli et al. [22] applied the SAA method to solve two chance constrained problems, namely, linear portfolio selection problem and blending problem with a joint chance constraint. Recently, Mancilla and Storer [17] formulated a stochastic scheduling problem using the SAA approach and proposed a heuristic method based on Benders decomposition.

Our main methodology builds upon the theory developed in Luedtke and Ahmed [16]. They first proposed to use the SAA approach to find feasible solutions and lower bounds on the optimal objective value of a general chance-constrained program. Keeping the same required risk level, they showed that the corresponding SAA counterpart yields a lower bound of the optimal objective value. To find a feasible solution, they showed that it suffices to solve a sample-based approximation with a smaller risk level. They also mathematically derived the required sample sizes in theory for having a lower bound or a feasible solution with high confidence. Our paper contributes to the literature by first employing the SAA approach to solve a class of production planning problems subject to a joint service-level constraint, which is typically computationally intensive.

1.2 Contributions of this paper

The main contributions of this paper are summarized as follows.

1. From the modeling perspective, we propose two new production planning models (with and without pricing decisions) subject to a joint service-level constraint. In classical stochastic production planning problems, unsatisfied demands are typically penalized by a linear backorder cost only. However, it is usually important for firms to maintain their reliability or credibility by persistently satisfying all the market demands in each period with high probability. Some firms use *α-service-level* (defined as the probability that the demand is fully satisfied) to measure their quality of service (QoS). This metric is yet neglected in most classical production planning models in the literature. This motivates us to incorporate a joint α-service-level constraint that ensures the market demands being satisfied in each period with a sufficiently high probability, so that the related firms can remain competitive and profitable.

Moreover, in most stochastic production planning models in the existing literature, the demand distributions in each period are given explicitly, while in real-life applications, it is usually difficult to deduce the true underlying demand distribution. The SAA reformulation can be done without knowing the exact demand distribution; however, a large amount of empirical data (more than 5000 samples of demands) is needed to solve the SAA reformulation under the nominal risk level and such amount of data may not be available in reality. We show that by using smaller risk parameter in the SAA reformulation, we are able to obtain good feasible solutions (within 5% of optimality) using much less empirical data (around 300 samples of demands), which makes the data-collection work less demanding. Also, our proposed models allow for nonstationary and generally correlated demands.

2. From the computational perspective, we employ the SAA method for chance-constrained programming and reformulate the two production planning models as MILPs using finite samples. However, due to the large amount of samples needed in the SAA reformulation, solving the resulting MILPs exactly is computationally intensive. Tuning the risk parameter in the SAA reformulation smaller than the required service level (i.e., more conservative), we achieve feasible solutions by solving the resulting MILPs via much fewer samples. The feasible solutions provide an upper (lower) bound on the cost-minimization (profit-maximization) problem. On the other hand, we also compute a lower (upper) bound on the cost-minimization (profit-maximization) problem by setting the risk parameter to be at least equal to the service level and solving multiple SAA counterparts with fewer number of samples.

3. A comprehensive numerical study has been conducted using three popular demand models with different patterns of demand correlations among periods (i.e., identical and independent demand distributions, Markov modulated demand process (MMDP) and autoregressive demands (AR models)), which are extensively used in theory and practice. For each problem instance with different demand model and different required service levels, both upper and lower bounds are computed and validated. We then compare our bounds with the optimal solutions (for reasonable problem sizes). It can be observed that the more samples we use, the less gap it has from our bounds to the optimal solutions. Also, the actual sample size needed to achieve at a given confidence level for both upper and lower bound solutions is a magnitude smaller than the theoretic bounds proposed in Luedtke and Ahmed [16]. We conduct sensitivity analysis for production planning with pricing and our numerical results show that when the demand is less sensitive to the price, the firm tends to increase the price while keeping the demand at the same level, in order to obtain a better profit.

1.3 Organization of the paper

The remainder of this paper is organized as follows. In Section 2, we introduce the notation and formulate the joint service-level constrained stochastic production planning problem. Section 3 formulates the production planning problem with pricing options. The computational results and insights for both problems with/without pricing options are given in Section 4. Finally, Section 5 concludes the paper and gives future research directions.

Throughout the paper, for notational convenience, we use a capital letter and its lowercase form to distinguish between a random variable and its realization. We use $\triangleq$ to indicate "is defined as", and $\mathbb{1}(A)$ is the indicator function taking value 1 if statement A is true and 0 otherwise. For any $x \in \mathbb{R}$, we denote $x^+ = \max\{x, 0\}$. We also use $\lceil x \rceil$ to denote the smallest integer that is no less than x and use $\lfloor x \rfloor$ to denote the largest integer not greater than x.

2 Production Planning with a Joint Service-Level Constraint

2.1 Mathematical formulation

Consider a finite horizon of T periods. The classic production planning problem decides the production quantities for each period simultaneously at the beginning of the planning horizon (denoted as $q_1, q_2, \ldots, q_T$). During each period t ($t = 1, \ldots, T$), demands are realized and three types of cost are incurred: production cost (with a per-unit cost $c_t > 0$), holding cost for on-hand inventories from period t to $t+1$ (with a per-unit cost $h_t > 0$), and penalty cost for backlogged demand (with a per-unit cost $p_t > 0$). The objective is to minimize the total cost over the T periods. Let $D_1, \ldots, D_T$ denote the random demands over the T periods and they may be independently distributed or correlated.

Let X_t and B_t be random variables that denote on-hand inventories and backorders at the end of period $t = 1, 2, \ldots, T$, respectively. Clearly, both X_t and B_t must be nonnegative. The initial inventory and backorder levels are denoted by x_0 and b_0, respectively. We formulate the production planning problem under a joint service-level constraint as

$$\begin{aligned}
\text{(PP)} \quad \min \quad & \sum_{t=1}^{T} (c_t q_t + h_t \mathbb{E}[X_t] + p_t \mathbb{E}[B_t]) && (1)\\
\text{s.t.} \quad & X_{t-1} + q_t + B_t = D_t + X_t + B_{t-1}, \quad \forall t = 1, \ldots, T, && (2)\\
& \mathbb{P}(X_t - B_t \geqslant 0, \forall t = 1, \ldots, T) \geqslant 1 - \theta, && (3)\\
& q_t \geqslant 0, \quad \forall t = 1, \ldots, T. && (4)
\end{aligned}$$

The objective (1) minimizes the total ordering cost, expected inventory cost and expected backlogging cost. In each period t, the incoming items are X_{t-1}, q_t and B_t while the outgoing items are X_t, B_{t-1} and D_t. To balance them, we formulate (2) as the *flow-balance constraints*. Constraint (3) requires that the probability of satisfying the demands in all T periods is at least $1-\theta$, which defines the service level.

2.2 Reformulation using the SAA approach

Consider N samples of demands over T periods denoted by $d^{(i)} = (d_1^{(i)}, \dots, d_T^{(i)})$ ($i = 1, 2, \dots, N$) where each sample i is equally likely to occur with probability $1/N$. The on-hand inventories and backorders vary according to demand samples, denoted by $x^{(i)} = (x_0^{(i)}, \dots, x_T^{(i)})$ and $b^{(i)} = (b_0^{(i)}, \dots, b_T^{(i)})$, respectively. The initial on-hand inventory and backorder are pre-determined regardless of the realization of random demands, i.e., $x_0^{(i)} = x_0$ and $b_0^{(i)} = b_0$ for all $i = 1, 2, \dots, N$. The ordering quantities $q_1, \dots q_t$ are decided before knowing the demand realizations, and thus do not depend on the specific samples.

In each sample i, the balance constraint (2) is presented as

$$x_{t-1}^{(i)} - x_t^{(i)} - b_{t-1}^{(i)} + b_t^{(i)} + q_t = d_t^{(i)}, \quad \forall t = 1, \dots, T. \tag{5}$$

We compute the total expected cost as:

$$\sum_{t=1}^{T} c_t q_t + \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \left(h_t x_t^{(i)} + p_t b_t^{(i)} \right).$$

The joint service-level constraint (3) is equivalent to

$$\sum_{i=1}^{N} \mathbb{1} \left\{ x_t^{(i)} \geqslant b_t^{(i)}, \forall t = 1, 2, \dots, T \right\} \geqslant \lceil (1-\theta) N \rceil. \tag{6}$$

Define binary variables $y^{(i)}$ such that $y^{(i)} = 1$ if and only if the i-th scenario is violated. We then replace the joint service-level constraint (6) by:

$$\begin{cases} x_t^{(i)} - b_t^{(i)} \geqslant -M_t^{(i)} y^{(i)}, \quad \forall t = 1, \dots, T, & (7) \\ \sum_{i=1}^{N} y^{(i)} \leqslant \lfloor \theta N \rfloor, & (8) \\ y \in \{0, 1\}^N. & (9) \end{cases}$$

Note that for each period t and sample i, the big-M coefficient $M_t^{(i)} = -x_0 + b_0 + \sum_{s=1}^{t} d_s^{(i)}$ is a valid upper bound for $-x_t^{(i)} + b_t^{(i)}$, because

$$\begin{aligned} b_t^{(i)} - x_t^{(i)} &= (d_t^{(i)} - q_t) + (b_{t-1}^{(i)} - x_{t-1}^{(i)}) \\ &= \sum_{s=1}^{t} (d_s^{(i)} - q_s) + (b_0 - x_0) \\ &\leqslant (b_0 - x_0) + \sum_{s=1}^{t} d_s^{(i)}. \end{aligned}$$

When $y^{(i)} = 0$, the constraint $x_t^{(i)} - b_t^{(i)} \geqslant 0$ is enforced for each $t = 1, 2, \dots, T$. When $y^{(i)} = 1$, the joint service-level constraint in the i-th sample can be violated and the total number of violated samples is no more than $\lfloor \theta N \rfloor$, ensured by the constraint $\sum_{i=1}^N y^{(i)} \leqslant \lfloor \theta N \rfloor$.

Therefore, we approximate a multi-period plan of optimal ordering quantities by using the following MILP model.

$$
\begin{aligned}
\text{(SAA-PP)} \quad \min \quad & \left\{ \sum_{t=1}^{T} c_t q_t + \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \left(h_t x_t^{(i)} + p_t b_t^{(i)} \right) \right\} \\
\text{s.t.} \quad & (5), (7)\text{–}(9), \\
& x_0^{(i)} = x_0,\ b_0^{(i)} = b_0, \quad \forall i = 1, \dots, N, && (10) \\
& x_t^{(i)}, b_t^{(i)}, q_t \geqslant 0, \quad \forall t = 1, \dots, T,\ i = 1, \dots, N. && (11)
\end{aligned}
$$

We present a necessary condition for any optimal production plan in the following proposition.

Proposition 1. For any optimal solution $\{q_t^*, x_t^{(i)*}, b_t^{(i)*}\}_{i,t}$ to the (SAA-PP), we have either $x_t^{(i)*} = 0$ or $b_t^{(i)*} = 0$ holds for all $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$.

Proof. We show the results by contradiction. Suppose that there exists an optimal solution with $x_t^{(i)} > 0$ and $b_t^{(i)} > 0$ for some $t \in \{1, 2, \dots, T\}$ and $i \in \{1, 2, \dots, N\}$. We can replace $x_t^{(i)}$ and $b_t^{(i)}$ by $\tilde{x}_t^{(i)} = x_t^{(i)} - \min\{x_t^{(i)}, b_t^{(i)}\}$ and $\tilde{b}_t^{(i)} = b_t^{(i)} - \min\{x_t^{(i)}, b_t^{(i)}\}$ while keeping the other decisions the same. Since the inventory level remains unchanged, i.e., $\tilde{x}_t^{(i)} - \tilde{b}_t^{(i)} = x_t^{(i)} - b_t^{(i)}$, all the constraints are satisfied under the new solution. Meanwhile, this new feasible solution decreases the total cost by $\frac{1}{N}(h_t + p_t) \cdot \min(x_t^{(i)}, b_t^{(i)}) > 0$. This contradicts with the fact that the original solution is optimal and thus completes the proof. □

The above proposition is also true when no service-level requirement is present [see, e.g., 26]. It asserts that even in the case of having a joint service-level constraint, there is no incentive to have backorders while holding positive inventories. This result is also true when there is no service-level requirement.

3 Model Variant with Pricing Options

3.1 Notation and mathematical formulation

We consider pricing decisions in the above production planning problem. In this variant, besides the ordering quantity q_t for each period $t = 1, 2, \dots, T$, the manager also decides the price r_t for each period $t = 1, 2, \dots, T$ at the beginning of the whole time horizon. The price r_t set by the manager affects the underlying demand distribution $D(t)$ and thus the realization d_t. The goal is to maximize the total expected profit over the T periods.

We interpret the random demand by a deterministic linear function in price r_t plus a noise term, i.e., $D_t(r_t) = -a_t r_t + \beta_t + \tilde{\epsilon}_t$, where $\tilde{\epsilon}_t$ is a random variable with $\mathbb{E}[\tilde{\epsilon}_t] = 0$ for $t = 1, \ldots, T$, and both $a_t, \beta_t > 0$. This demand model is well known as the additive demand model in the literature (see, e.g., Mills [19]). It allows for correlated demands over periods, which indicates that $\{\tilde{\epsilon}_t\}_{t=1}^T$ are not necessarily independent random variables.

We further assume that once the demand is realized in period t, it establishes a contract between the buyer and the retailer with a unit price of r_t. In other words, given that the realized demand $D_t = d_t$, it immediately incurs a revenue of $r_t d_t$, no matter when the demand is satisfied. The price $r = (r_1, r_2, \ldots, r_T)^\mathsf{T}$ is chosen from a given set $P \subseteq \mathbb{R}^T$. We can then formulate the production planning problem with pricing options as

$$\begin{aligned}
\text{(PO)} \quad \max \quad & \sum_{t=1}^{T} \Big(r_t \mathbb{E}[D_t] - c_t q_t - h_t \mathbb{E}[X_t] - p_t \mathbb{E}[B_t] \Big) \\
\text{s.t.} \quad & (2)\text{–}(4), \\
& D_t = -a_t r_t + \beta_t + \tilde{\epsilon}_t, \quad \forall t = 1, \ldots, T, \qquad (12) \\
& r \in P.
\end{aligned}$$

In the formulation above, the constraints (2)–(4) are carried from the model without pricing options. The constraint (12) shows the relationship between price and demand in each period.

3.2 Sample average approximation reformulation

We reformulate the (PO) model using N sample data, namely $d^{(i)} = (d_1^{(i)}, d_2^{(i)}, \ldots, d_T^{(i)})^\mathsf{T}$, $i = 1, 2, \ldots, N$. For the i-th sample, we use $\epsilon_t^{(i)}$ to denote the realizations of $\tilde{\epsilon}_t$ and other notations remain the same. Then the total expected profit is given by

$$-\sum_{t=1}^{T} c_t q_t + \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \Big[r_t d_t^{(i)} - h_t x_t^{(i)} - p_t b_t^{(i)} \Big].$$

Hence, using the linear relationship between demand and price, i.e., $d_t^{(i)} = -a_t r_t + \beta_t + \epsilon_t^{(i)}$, we reformulate the (PO) model as:

(SAA-PO)

$$\begin{aligned}
\max \quad & \sum_{t=1}^{T} \big(-c_t q_t - a_t r_t^2 + \beta_t r_t \big) + \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \big(\epsilon_t^{(i)} r_t - h_t x_t^{(i)} - p_t b_t^{(i)} \big) \\
\text{s.t.} \quad & x_{t-1}^{(i)} - x_t^{(i)} - b_{t-1}^{(i)} + b_t^{(i)} + q_t = -a_t r_t + \beta_t + \epsilon_t^{(i)}, \forall t = 1, \ldots, T,\ i = 1, \ldots, N, \\
& x_t^{(i)} - b_t^{(i)} \geqslant -\Big(-x_0 + b_0 + \sum_{s=1}^{t} \big(-a_s r_s + \beta_s + \epsilon_s^{(i)} \big) \Big) y^{(i)},
\end{aligned}$$

$$\forall t = 1, \ldots, T,\ i = 1, \ldots, N, \tag{13}$$

(8)–(11),

$r \in P$.

Here, the probabilistic constraint (3) in the (PO) model is equivalent to constraints (13), (8) and (9), for which we define new binary variables $y^{(i)}$ ($i = 1, \ldots, N$). However, the above model still involve nonlinear terms r_t^2 in the objective function and bilinear terms $r_s y^{(i)}$ in constraint (13). We reformulate it as a linear model for two specific price sets with either discrete or continuous price choices, where price decisions p_t, $t = 1, ..., T$, are independently determined for each period. For the former, the price is drawn from a set of finite possible prices, denoted by set $R_t = \{\gamma_1^t, \ldots, \gamma_{m_t}^t\}$. In this case, the price set can be written as $P = \bigtimes_{t=1}^T R_t$. For the latter, we consider the possible price r_t chosen from a price interval $[L_t, U_t]$ given for each period t and the price set is specified as $P = \bigtimes_{t=1}^T [L_t, U_t]$. We will give MILP models for each price set in the following subsections. Note that our model is also capable of describing the relationship among the prices in each period. For example, if the prices are not allowed to increase from periods to periods, we can simply add the constraint $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_T$ to our model while the complexity of the resulting model remains the same.

3.2.1 Discrete price set

Consider a finite set $R_t = \{\gamma_1^t, \ldots, \gamma_{m_t}^t\}$ from which product price is chosen in each period. Define a binary decision variable u_{jt} to indicate whether using the j-th price option, i.e., γ_j^t ($j = 1, \ldots, m_t$). In each period, $u_{jt} = 1$ if $r_t = \gamma_j^t$ and $u_{jt} = 0$ otherwise. To ensure only one price is used from the set $(\gamma_1^t, \gamma_2^t, \ldots, \gamma_{m_t}^t)$ in each period t, we require $\sum_{j=1}^{m_t} u_{jt} = 1$ for each t. Then, the quadratic term r_t^2 in objective function can be expressed by a linear term:

$$r_t^2 = \sum_{j=1}^{m_t} (\gamma_j^t)^2 u_{jt}. \tag{14}$$

Similarly, for any $t = 1, \ldots, T$ and $i = 1, \ldots, N$, the nonlinear term $r_t y^{(i)}$ in the joint service-level constraint can be rewritten as

$$r_t y^{(i)} = \sum_{j=1}^{m_t} \gamma_j^t u_{jt} y^{(i)}. \tag{15}$$

To further linearize the term $u_{jt} y^{(i)}$ in (15), we introduce another decision variable $v_{jt}^{(i)}$ to replace $u_{jt} y^{(i)}$ and add the following McCormick inequalities to force $v_{jt}^{(i)} = u_{jt} y^{(i)}$ (see

McCormick [18]):

$$\begin{cases} v_{jt}^{(i)} \leqslant u_{jt} \\ v_{jt}^{(i)} \leqslant y^{(i)} \\ v_{jt}^{(i)} \geqslant u_{jt} + y^{(i)} - 1 \\ v_{jt}^{(i)} \geqslant 0. \end{cases} \tag{16}$$

Here both u_{jt} and $y^{(i)}$ are binary variables. If $u_{jt} = 1$ and $y^{(i)} = 1$, then $v_{jt}^{(i)} = 1$; otherwise, the constraints enforce $v_{jt}^{(i)} = 0$. Therefore, for discrete pricing options, using equalities (14) and (15), we can formulate the following (PP-D) model:

$$\begin{aligned}
&\text{(PP-D)} \\
&\max \quad \sum_{t=1}^{T}\Big(-c_t q_t - a_t \sum_{j=1}^{m_t} (\gamma_j^t)^2 u_{jt} + \beta_t \sum_{j=1}^{m_t} \gamma_j^t u_{jt}\Big) + \frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N}\Big(\epsilon_t^{(i)} \sum_{j=1}^{m_t} \gamma_j^t u_{jt} - h_t x_t^{(i)} - p_t b_t^{(i)}\Big) \\
&\text{s.t.} \quad x_{t-1}^{(i)} - x_t^{(i)} - b_{t-1}^{(i)} + b_t^{(i)} + q_t = -a_t \sum_{j=1}^{m_t} \gamma_j^t u_{jt} + \beta_t + \epsilon_t^{(i)}, \forall t = 1, \dots, T,\ i = 1, \dots, N, \\
&\qquad x_t^{(i)} - b_t^{(i)} \geqslant (x_0 - b_0) y^{(i)} + \sum_{s=1}^{t}\Big(a_s \sum_{j=1}^{m_s} \gamma_j^s v_{js}^{(i)} - \beta_s y^{(i)} - \epsilon_s^{(i)} y^{(i)}\Big), \\
&\qquad\qquad \forall t = 1, \dots, T,\ i = 1, \dots, N, \\
&\qquad (8)\text{–}(11), \\
&\qquad (16), \quad \forall t = 1, \dots, T,\ i = 1, \dots, N,\ j = 1, \dots, m_t, \\
&\qquad \sum_{j=1}^{m_t} u_{jt} = 1, \quad \forall t = 1, \dots, T, \\
&\qquad u_{jt} \in \{0, 1\}, v_{jt}^{(i)} \geqslant 0, \quad \forall t = 1, \dots, T,\ i = 1, \dots, N,\ j = 1, \dots, m_t.
\end{aligned}$$

3.2.2 Continuous price set

For continuous pricing options, the price in period t is chosen from the set $P_t = [L_t, R_t]$. We first note that the price r_t must be non-negative, hence $L_t \geqslant 0$. Also, r_t should be bounded above by β_t/a_t; otherwise the expected demand $-a_t r_t + \beta_t < 0$, which is unlikely to happen in reality. Hence, the retailer will never set a price higher than β_t/a_t, and therefore, $U_t \leqslant \beta_t/a_t$.

For the nonlinear term $r_t y^{(i)}$ presented in the joint service-level constraint, we linearize it by introducing a new decision variable $w_t^{(i)}$. The following sets of linear inequalities enforce $w_t^{(i)} = r_t y^{(i)}$ when $y^{(i)}$ is a binary:

$$\begin{cases} w_t^{(i)} \leqslant r_t \\ w_t^{(i)} \leqslant U_t y^{(i)} \\ w_t^{(i)} \geqslant r_t + U_t(y^{(i)} - 1) \\ w_t^{(i)} \geqslant 0. \end{cases} \tag{17}$$

Finally we can formulate the following (PP-C) model with quadratic objective for production planning problem with continuous pricing options as follows

(PP-C)

$$
\begin{aligned}
\max \quad & \sum_{t=1}^{T} \big(-c_t q_t - a_t r_t^2 + \beta_t r_t \big) + \frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} \big(\epsilon_t^{(i)} r_t - h_t x_t^{(i)} - p_t b_t^{(i)} \big) \\
\text{s.t.} \quad & x_{t-1}^{(i)} - x_t^{(i)} - b_{t-1}^{(i)} + b_t^{(i)} + q_t = -a_t r_t + \beta_t + \epsilon_t^{(i)}, \forall t = 1, \ldots, T,\ i = 1, \ldots, N, \\
& x_t^{(i)} - b_t^{(i)} \geqslant (x_0 - b_0) y^{(i)} + \sum_{s=1}^{t} \big(a_s w_s^{(i)} - \beta_s y^{(i)} - \epsilon_s^{(i)} y^{(i)} \big), \forall t = 1, \ldots, T,\ i = 1, \ldots, N, \\
& (8)\text{–}(11), \\
& (17), \quad \forall t = 1, \ldots, T,\ i = 1, \ldots, N, \\
& y^{(i)} \in \{0, 1\}, \quad \forall i = 1, \ldots, N, \\
& L_t \leqslant r_t \leqslant U_t, \quad \forall t = 1, \ldots, T.
\end{aligned}
$$

4 Computational Results

4.1 Solution methods

In general, an MILP reformulation of a chance-constrained program is computationally intractable since it usually requires a large number of Monte Carlo samples to attain solution accuracy. Luedtke and Ahmed [16] suggested using the SAA approach for solving general chance-constrained programs, and derived theocratical sample-size bounds for obtaining solutions that satisfy the chance constraints with certain confidence for specific risk levels. Specifically, consider a generic chance-constrained program:

$$(\mathrm{P}_\theta): \quad z_\theta^* = \min\{f(x) : x \in X_\theta\},$$

where $X_\theta = \Big\{ x \in X : \mathbb{P}\left\{ G(x, \xi) \geqslant \mathbf{0} \right\} \geqslant 1 - \theta \Big\}$. Here $X \subseteq \mathbb{R}^n$ represents a deterministic feasible region (i.e., given by the constraints (5), (10), (11)), $f : \mathbb{R}^n \to \mathbb{R}$ represents the objective to be minimized, ξ is a random vector with support $\Xi \subseteq \mathbb{R}^d$, $G : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is a given constraint mapping and θ is a risk parameter of service level. We assume that z_θ^* exists and is finite.

The SAA counterpart of the chance-constrained problem (P_θ) with risk parameter α is defined as

$$(\mathrm{P}_\alpha^N): \quad \hat{z}_\alpha^N = \min\{f(x) : x \in X_\alpha\},$$

where $X_\alpha = \Big\{ x \in X : \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}\Big(G(x, \xi^i) \geqslant \mathbf{0} \Big) \geqslant 1 - \alpha \Big\}$.

Feasible solutions: To obtain feasible solutions of (P_θ), one can choose a smaller risk parameter $\alpha < \theta$ and solve the SAA counterpart (P_α^N). As shown in Luedtke and Ahmed

[16], if the sample size

$$N \geqslant \frac{1}{2(\theta-\alpha)^2}\log\left(\frac{|X\backslash X_\theta|}{\delta}\right), \tag{18}$$

then solving (P_α^N) will yield a feasible solution to (P_θ) with probability at least $1-\delta$. This gives a theoretical sample size to guarantee a feasible solution for P_θ using the SAA approach with a confidence level $1-\delta$.

Note that the exact value of $|X\backslash X_\theta|$ is often not available, and we can estimate the least value of the right-hand side in (18) by setting $|X\backslash X_\theta| = 1$. For example, When $\theta = 0.02$ and $\alpha = 0$, to guarantee a feasible solution with probability at least 90%, we need at least $N \geqslant \frac{1}{2\times 0.02^2}\log(1/0.1) \approx 2878$ number of samples. Similarly, when $\theta = 0.05$ and $\alpha = 0$, to guarantee a feasible solution with probability at least 90%, we need at least $N \geqslant \frac{1}{2\times 0.05^2}\log(1/0.1) \approx 460$ number of samples. In our later computational studies, we test different values of N that are much smaller than the suggested theoretical sample sizes, and demonstrate the performance of the resulting solutions in out-of-sample simulation tests.

Lower bounds: To obtain lower bounds on the original optimization problem (P_θ), we set $\alpha = \theta$ and to generate M SAA problems, namely, ($\mathrm{P}_{\theta,i}^N$) ($i = 1, 2, \ldots, M$). Then we solve each sample-based problem and obtain a set of optimal objective values, denoted by $\hat{z}_{\theta,i}^N$ ($i = 1, 2, \ldots, M$). The L-th minimum value among all M optimal objective value is denoted by $\hat{z}_{\theta,[L]}^N$. Then, according to Luedtke and Ahmed [16], the following result holds:

$$\mathbb{P}(\hat{z}_{\theta,[L]}^N \leqslant z_\theta^*) \geqslant 1 - \sum_{i=0}^{L-1}\binom{M}{i}(1/2)^M \tag{19}$$

for large enough N relative to ϵ (e.g., $N\epsilon \geqslant 10$). Hence, we can say that $\hat{z}_{\theta,[L]}^N$ is a lower bound of the objective value with a confidence level $1 - \sum_{i=0}^{L-1}\binom{M}{i}(1/2)^M$.

We test the effectiveness of the SAA approach applied to both models in this paper with and without the pricing option. We use CPLEX 12.5.1 for solving all MILP models. All the computations are performed on a 3.40GHz Intel(R) Xeon(R) CPU.

4.2 Stochastic production planning problem

We present numerical results on the stochastic production planning problem. We numerically solve the appropriate SAA counterpart problems and compute both the upper bounds and the lower bounds for optimal objective values. We also compute the required sample size in practice to show the effectiveness of our approach.

4.2.1 Parameter setting

We use randomly generated instances based on different demand distributions to demonstrate the general features of our models and approaches. We test both i.i.d. demand and correlated demand. We consider the total number of periods $T = 5$, and assume stationary unit ordering cost, unit holding cost, and unit backlogging cost in each period, which are $c = 5$, $h = 1$, and $p = 10$, respectively.

The i.i.d demand in each period follows Poisson distribution with mean value 20. For the correlated demand, we consider both Markov Modulated Demand Process (MMDP) (see, e.g., Chen and Song [5]) and Autoregressive Model of Order 1 (AR(1)) (see, e.g., Mills [20]).

The demands generated from MMDP have three states corresponding to the *state of economy*: poor (1), fair (2), and good (3). In each period t, given that the current state is $i_t \in \{1, 2, 3\}$, we test cases where the demand distributions are Poisson with mean value $10i_t$. We also assume that the state of the economy follows a Markov chain with the initial state 1 (i.e., poor) and the transition probability matrix

$$\mathcal{P} = \begin{pmatrix} 0.2 & 0.5 & 0.3 \\ 0.4 & 0.2 & 0.4 \\ 0.1 & 0.6 & 0.3 \end{pmatrix}.$$

For the AR(1) demand case, the demands in period t satisfy $d_t = d_{t-1} + \eta$, where the noise term η is normally distributed with mean 0 and standard deviation 1. We set the initial demand $d_0 = 20$. In all our computations, we test the problems with required service-level $\theta = 0.02$ and 0.05.

4.2.2 Feasible solutions

We aim at demonstrating the effectiveness of the SAA approach for finding feasible solutions. To compute for feasible solutions, we set the risk level $\alpha = 0$. This gives us a more conservative SAA counterpart problem. However, a relatively small sample size N is needed to compute feasible solutions. Also, we numerically compute the solution to the SAA counterpart which uses the required service level as the risk parameter (i.e., $\alpha = \theta$). We compare the statistics of using these two different values for the risk level α.

Our numerical test consists of two parts. First, we generate N samples and solve the corresponding SAA instances each time. The above process is repeated $M = 10$ times so that we obtain 10 solutions for the same problem. Second, to validate if all these 10 solutions are feasible, we conduct a *posteriori* check to compute the risk for each solution: we generate a simulation sample with $N' = 10,000$ scenarios, and check the number of scenarios that are violated under the larger problem for each given solution. The solution risk is then calculated as

$$\mathcal{R} = \frac{\text{number of violated scenarios}}{N'}.$$

If $\mathcal{R} < \theta$, the service-level requirements are satisfied; otherwise, the solution is not feasible. For solution risk, we report the average (Avg), minimum (Min), maximum (Max), and sample standard deviation (σ) over the solutions given by the 10 SAA problems. We also report the number of feasible solutions as well as the average, minimum, maximum, and sample standard deviation of the cost over these feasible solutions.

Tables 1–7 summarize our computational results for finding feasible solutions to the stochastic production planning problem. When not applicable, we indicate *** in the corresponding entry of each table. Our observations are summarized as follows:

1. From each table, we observe that as the sample size N grows, the average solution risk decreases and the number of feasible solutions increases. This is because as more samples are used, more constraints are being enforced into the model, which leads to a smaller feasible region. Hence, as the sample size grows, we can obtain more conservative solutions by solving the SAA problems which have lower solution risks and a higher likelihood to be feasible at the nominal risk level θ.

2. We observe that using $\alpha = 0$ requires much fewer samples to achieve a feasible solution than using $\alpha = \theta$. For example, in Table 1, we only need 300 samples to get a feasible solution with confidence level 90% by using $\alpha = 0$. However, solving the SAA reformulation at the nominal risk level $\alpha = 0.02$ requires at least 3000 samples to have a confidence level of 80%. Note that the problem size grows as the number of samples increases, we conclude that solving the SAA problem by setting risk level $\alpha = 0$ is more efficient than solving the original SAA problem at the nominal risk level in terms of obtaining feasible solutions.

3. We also notice that the required sample sizes in our tests are also smaller than the theoretical bound given in Luedtke and Ahmed [16]. For instance, to achieve 90% confidence level, the theoretical required sample sizes calculated by (18) in Section 4.1 are $\bar{N}_{0.02} = 2878$ and $\bar{N}_{0.05} = 460$ for $\theta = 0.02$ and $\theta = 0.05$, respectively; however, from Tables 3 and 4, we can see that the actual sample sizes in practice are $N_{0.02} = 300$ and $N_{0.05} = 100$, respectively. The smaller sample size not only makes the computation more efficient, but also makes the data-collection work less demanding.

4. In terms of the costs for feasible solutions, we observe that using $\alpha = 0$ yields a higher average cost and a higher variance among all feasible solutions than using $\alpha = \theta$. For example, in Table 6 the average cost for feasible solutions is 640.7 using $\alpha = 0$ and $N = 50$ samples, as compared to the average cost of 605.58 by using $\alpha = \theta$ and $N = 3000$. Hence, although using a smaller risk level $\alpha = 0$ is more efficient to compute a feasible solution under a given confidence level, it might yield a more conservative solution that has a higher cost than solving the SAA problems under the nominal risk level because no scenarios can be violated when risk level α is set to be 0. Therefore,

using a smaller risk level $\alpha = 0$ produces a feasible solution and an upper bound on the objective values efficiently.

5. Another observation from the tables is that the average feasible solutions cost and their standard deviations are more erratic when we use a risk level $\alpha = 0$, as compared to $\alpha = \theta$. The reason is that when using a smaller risk level $\alpha = 0$, the feasible solutions cost are more sensitive with respect to the sample size. This also explains why a smaller sample size is required when we have a smaller risk level.

Table 1: Solution results of i.i.d. demand for stochastic production planning problem without pricing for $\theta = 0.02$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	50	0.054	0.003	0.146	0.043	1	782.64	782.64	782.64	***
	100	0.025	0.003	0.054	0.015	4	764.37	715.65	818.83	40.12
	200	0.016	0.004	0.033	0.008	8	752.89	724.38	788.75	22.23
	300	0.011	0.005	0.021	0.005	9	771.58	744.26	808.36	19.02
	400	0.007	0.003	0.012	0.002	10	771.47	742.12	804.69	17.89
0.02	250	0.033	0.014	0.046	0.153	1	725.75	725.75	725.75	***
	500	0.028	0.017	0.044	0.152	3	713.58	709.00	716.47	3.28
	1000	0.025	0.016	0.035	0.132	3	715.34	714.16	716.43	0.93
	2000	0.021	0.017	0.029	0.004	4	712.39	710.30	718.24	3.38
	3000	0.019	0.017	0.024	0.002	8	711.26	707.06	714.28	2.36

Table 2: Solution results of i.i.d. demand for stochastic production planning problem without pricing for $\theta = 0.05$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	20	0.113	0.017	0.287	0.076	2	765.78	721.8	809.75	43.97
	60	0.051	0.008	0.084	0.024	4	714.18	691.37	740.97	18.71
	100	0.025	0.003	0.054	0.015	9	743.15	707.62	818.83	35.37
	140	0.021	0.003	0.041	0.011	10	733.92	698.41	790.60	25.23
0.05	500	0.064	0.048	0.080	0.010	1	671.83	671.83	671.83	***
	750	0.055	0.041	0.066	0.008	3	675.08	673.71	677.44	1.67
	1000	0.054	0.047	0.066	0.006	3	675.72	674.03	677.61	1.47
	2000	0.048	0.039	0.055	0.004	8	675.16	671.48	680.13	2.66

Table 3: Solution results of MMDP for stochastic production planning problem without pricing decisions for $\theta = 0.02$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	50	0.051	0.008	0.118	0.033	2	990.17	946.06	1034.28	44.11
	100	0.033	0.012	0.065	0.016	2	897.52	895.45	899.58	2.07
	200	0.019	0.003	0.036	0.011	5	919.38	846.43	980.29	45.91
	300	0.012	0.004	0.027	0.006	9	898.96	862.15	940.07	28.33
	400	0.006	0.001	0.012	0.004	10	944.01	896.75	1016.14	40.52
0.05	500	0.031	0.016	0.044	0.008	1	845.08	845.08	845.08	***
	1000	0.025	0.017	0.032	0.005	3	842.58	839.77	847.47	3.48
	2000	0.021	0.016	0.027	0.003	4	843.95	839.18	849.02	3.76
	3000	0.020	0.017	0.024	0.002	3	844.72	842.03	847.13	2.09

Table 4: Solution results of MMDP for stochastic production planning problem without pricing decisions for $\theta = 0.05$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	25	0.124	0.039	0.274	0.089	2	912.14	877.28	947.00	34.86
	50	0.064	0.025	0.129	0.034	4	853.81	828.44	885.00	22.60
	75	0.036	0.009	0.090	0.025	8	862.35	806.27	959.87	42.71
	100	0.033	0.012	0.065	0.016	9	857.77	807.34	899.58	35.42
	125	0.025	0.011	0.039	0.009	10	878.36	851.79	922.89	22.17
0.05	500	0.061	0.046	0.075	0.009	1	797.85	797.85	797.85	***
	1000	0.054	0.048	0.065	0.005	2	793.13	792.99	793.26	0.14
	2000	0.052	0.045	0.060	0.004	2	794.11	793.42	794.80	0.69

Table 5: Solution results of AR(1) demand for stochastic production planning problem without pricing decisions for $\theta = 0.02$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	25	0.057	0.012	0.094	0.025	1	640.96	640.96	640.96	***
	50	0.041	0.008	0.113	0.033	3	650.48	642.68	664.28	9.79
	75	0.023	0.010	0.048	0.012	6	647.59	638.01	665.27	10.11
	100	0.012	0.001	0.029	0.008	9	659.12	628.93	718.99	29.51
	125	0.010	0.001	0.019	0.006	10	659.67	630.07	720.3	28.09
0.05	500	0.025	0.015	0.042	0.008	2	632.46	630.9	634.02	***
	1000	0.02	0.013	0.026	0.005	5	631.43	626.43	634.18	2.94
	2000	0.018	0.016	0.022	0.002	8	627.75	626.10	633.12	2.17
	3000	0.018	0.016	0.022	0.002	7	627.94	626.14	629.87	1.03
	4000	0.018	0.016	0.02	0.001	10	627.31	625.81	628.49	0.82

Table 6: Solution results of AR(1) demand for stochastic production planning problem without pricing decisions for $\theta = 0.05$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	20	0.080	0.017	0.157	0.053	4	641.93	624.05	682.55	23.65
	30	0.044	0.007	0.094	0.031	6	633.26	616.77	676.87	20.89
	40	0.035	0.003	0.091	0.030	7	643.36	614.33	672.73	18.30
	50	0.039	0.002	0.094	0.033	7	640.70	613.06	674.98	22.12
	60	0.020	0.001	0.037	0.012	10	638.14	610.63	685.65	21.62
0.05	500	0.055	0.038	0.091	0.016	5	606.78	604.76	609.82	1.68
	1000	0.049	0.041	0.058	0.006	5	606.33	604.60	608.62	1.56
	2000	0.048	0.042	0.055	0.004	7	605.51	604.57	607.32	0.87
	3000	0.046	0.043	0.051	0.003	8	605.58	604.42	606.82	0.65
	4000	0.045	0.043	0.048	0.002	10	605.33	604.06	606.36	0.66

Table 7: Solution results of AR(1) demand for stochastic production planning problem without pricing decisions for $\theta = 0.08$

		Solution Risk				Feasible Solutions Cost				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	10	0.077	0.005	0.171	0.055	6	613.95	581.9	660.6	26.43
	20	0.080	0.017	0.157	0.052	4	641.93	624.05	682.55	23.65
	30	0.034	0.007	0.094	0.026	9	631.02	591.03	674.47	27.58
	40	0.034	0.003	0.102	0.031	9	639.23	611.13	671.20	20.33
	50	0.041	0.008	0.113	0.033	8	635.51	608.1	664.28	18.28
	60	0.027	0.008	0.036	0.008	10	627.14	610.87	659.42	13.02
	70	0.012	0.003	0.031	0.009	10	650.43	622.06	673.03	14.48
0.08	250	0.087	0.065	0.121	0.016	4	595.24	591.68	600.42	3.19
	500	0.076	0.059	0.086	0.010	6	595.00	591.12	599.02	2.95
	750	0.079	0.061	0.110	0.013	8	594.61	592.63	599.95	2.24
	1000	0.078	0.070	0.091	0.006	9	593.46	592.45	594.96	0.93
	1500	0.076	0.066	0.086	0.006	9	594.07	592.20	599.41	2.06

4.2.3 Lower bounds

To obtain the lower bounds for the stochastic production planning problem, we follow the same settings in Luedtke and Ahmed [16] by choosing $\alpha = \theta$ and $M = 10$. We then take the L-th minimum optimal objective value among all $M = 10$ runs. According to (19), the confidence levels of using $L = 1, \ldots, 4$ are 0.999, 0.989, 0.945 and 0.828, respectively.

In addition to the lower bounds computed at each confidence level, we also report optimality gaps, defined as the percentage that the lower bound is below the cost of best feasible solution (i.e., the minimum cost among all feasible solutions, given by Tables 1–6).

Tables 8–14 report the test results. Combining the test results on the lower bounds and the results of feasible solutions in Section 4.2.2, we can obtain the range of the optimal cost. For example, in the i.i.d. demand case with service level $\theta = 0.02$, solving $M = 10$ SAA instances with sample size $N = 250$ yields a feasible solution of cost 725.75 shown in Table 1 while getting a lower bound 675.24 with a confidence level 0.999 shown in Table 8. This means that we have at least 99.9% confidence to say that the optimal solution is at most $(725.75 - 675.24)/725.75 \times 100\% \approx 6.96\%$ less costly than the best feasible solution we get. Similarly, we can analyze the problem with other demand cases and different service level parameters using corresponding tables.

From these results, we observe that as sample size N becomes larger, the lower bound becomes larger and the gap becomes smaller at each confidence level. When the gap reaches zero, we come to a conclusion that the best feasible solution is the optimal solution with the corresponding confidence level. For example, as we notice from Table 13, when the sample size $N = 2000$, we have confidence at least 82.8% that the feasible solution of cost 604.57 is optimal; when the sample size raises to $N = 3000$, our confidence increases from 82.8% to 99.9%. Thus, for a certain confidence level, we can make a better estimation of the optimal solution of an SAA problem as we increase the sample size N.

Table 8: Lower bounds of i.i.d. demand for stochastic production planning problem without pricing for $\alpha = \theta = 0.02$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
250	675.24	683.10	683.35	684.52	6.96%	5.88%	5.84%	5.68%
500	675.83	691.38	691.94	692.90	4.68%	2.49%	2.41%	2.27%
1000	687.04	695.08	695.46	700.28	3.80%	2.67%	2.62%	1.94%
2000	694.73	698.35	699.20	702.01	2.19%	1.68%	1.56%	1.17%
3000	701.10	702.49	707.06	709.11	0.84%	0.65%	0.00%	−0.29%

Table 9: Lower bounds of i.i.d. demand for stochastic production planning problem without pricing for $\alpha = \theta = 0.05$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	655.94	655.97	658.63	660.89	2.36%	2.36%	1.96%	1.63%
750	662.20	663.93	664.02	667.83	1.71%	1.45%	1.44%	0.87%
1000	661.15	662.96	664.37	667.49	1.91%	1.64%	1.43%	0.97%
2000	668.37	669.04	671.48	673.31	0.46%	0.36%	0.00%	−0.27%

4.3 Production planning with pricing options

In this section, we report the computational results of multi-period joint service-level constrained production planning with pricing options. We also conduct sensitivity analysis for this model.

4.3.1 Parameter setting

Consider the continuous pricing in the test instances. The setting of cost parameters is the same as those in Section 4.2.1. Moreover, we set $a_t = -5$ and $\beta_t = 200$ in the function $d_t(r_t) = a_t r_t + \beta_t + \tilde{\epsilon}_t$ for all $t = 1, \ldots, T$. The noise term $\tilde{\epsilon}_t$ follows normal distribution

Table 10: Lower bounds of MMDP for stochastic production planning problem without pricing for $\alpha = \theta = 0.02$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	801.63	803.84	808.33	813.20	5.14%	4.88%	4.35%	3.77%
1000	812.40	813.36	817.94	824.20	3.26%	3.14%	2.60%	1.85%
2000	824.67	830.47	830.98	832.43	1.73%	1.04%	0.98%	0.80%
3000	830.88	835.12	836.84	837.06	1.32%	0.82%	0.62%	0.59%

Table 11: Lower bounds of MMDP for stochastic production planning problem without pricing for $\alpha = \theta = 0.05$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	768.09	770.16	770.58	774.45	3.73%	3.47%	3.42%	2.93%
1000	775.65	780.50	780.80	782.01	2.19%	1.57%	1.54%	1.38%
2000	778.77	780.37	784.50	784.76	1.85%	1.64%	1.12%	1.09%

Table 12: Lower bounds of AR(1) demand for stochastic production planning problem without pricing for $\alpha = \theta = 0.02$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	612.87	614.25	615.61	617.04	2.86%	2.64%	2.42%	2.20%
1000	620.22	620.25	620.81	621.43	0.99%	0.99%	0.90%	0.80%
2000	622.00	625.37	626.10	626.41	0.66%	0.12%	0.00%	−0.05%
3000	621.54	625.20	625.62	626.14	0.73%	0.15%	0.08%	0.00%
4000	625.81	626.44	626.92	626.95	0.00%	−0.10%	−0.18%	−0.18%

Table 13: Lower bounds of AR(1) demand for stochastic production planning problem without pricing for $\alpha = \theta = 0.05$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	584.34	597.29	600.79	600.91	3.38%	1.23%	0.66%	0.64%
1000	599.22	600.58	601.15	602.62	0.89%	0.66%	0.57%	0.33%
2000	600.56	602.23	603.20	604.57	0.66%	0.39%	0.23%	0.00%
3000	604.42	604.80	605.14	605.28	0.00%	−0.06%	−0.12%	−0.14%
4000	604.06	604.44	604.98	605.14	0.00%	−0.06%	−0.15%	−0.18%

Table 14: Lower bounds of AR(1) demand for stochastic production planning problem without pricing for $\alpha = \theta = 0.08$

	LB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
250	579.12	586.06	588.48	588.66	2.12%	0.95%	0.54%	0.51%
500	586.24	589.20	590.34	591.12	0.82%	0.33%	0.13%	0.00%
750	586.56	587.81	592.63	593.29	1.03%	0.81%	0.00%	-0.11%
1000	586.99	592.45	592.52	592.74	0.92%	0.00%	-0.01%	-0.05%
1500	591.51	592.20	592.30	592.72	0.12%	0.00%	-0.01%	-0.08%

with mean 0 and standard deviation 22 for all $t = 1, \ldots, T$. We also assume that the pricing range in each period t is between $W_t^L = 18$ and $W_t^U = 40$. We fix the required service level $\theta = 0.02$.

4.3.2 Feasible solutions

Table 15 reports statistics of the solutions of i.i.d. demand for production planning with pricing options. Table 16 reports statistics of the solutions of AR(1) demand for production planning with pricing options. The insights of our numerical results are summarized as follows:

1. As the sample size N grows, the average solution risk decreases and the number of feasible solutions increases since more constraints are being enforced into the model, which leads to a smaller feasible region. Hence, as the sample size increases, solving the SAA counterpart under any fixed risk level yields a lower solution risk and a higher likelihood to be feasible at the nominal risk level θ.

2. From Table 15 and Table 16, we observe that using $\alpha = 0$ requires much fewer samples to achieve a feasible solution than using $\alpha = \theta$. For example, in Table 15, we generate 300 samples to obtain a feasible solution with confidence level 100% by using $\alpha = 0$. However, even using 2500 samples in the SAA reformulation, with risk level $\alpha = 0.02$, can only find a feasible solution at a confidence level of 60%. In our numerical test, solving an SAA reformulation that involves more than 2500 samples is computationally intractable (more than three CPU minutes for each instance). Therefore, an efficient way to compute a feasible solution is to solve the SAA reformulation with a more conservative risk level $\alpha = 0$. The smaller sample size not only makes the computation more efficient, but also makes the data-collection work less demanding.

3. In terms of the profit for feasible solutions, we observe that using $\alpha = 0$ yields a lower average profit and a higher variance among all feasible solutions than using $\alpha = 0.02$. For example, in Table 16 the average profit for feasible solution is 6365.02 using $\alpha = 0$ and $N = 200$ samples, as compared to the average profit of 6567.83 by setting $\alpha = 0.02$ and $N = 2000$. Hence, although using a smaller risk level $\alpha = 0$ is more efficient to compute a feasible solution under a given confidence level, it will yield a more conservative solution that has a lower profit than solving the SAA problems under the nominal risk level. Therefore, using a smaller risk level $\alpha = 0$ helps us to find a feasible solution and a lower bound on the total profit efficiently.

Table 15: Solution results of i.i.d. demand for production planning with pricing for $\theta = 0.02$

		Solution Risk				Profit for Feasible Solutions				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0.00	50	0.047	0.020	0.148	0.036	1	6652.42	6652.42	6652.42	***
	100	0.024	0.008	0.045	0.010	4	6479.34	6308.48	6581.04	104.64
	200	0.017	0.008	0.027	0.006	8	6534.51	6362.91	6678.65	100.52
	300	0.010	0.005	0.020	0.004	10	6438.01	6261.56	6647.29	119.30
0.02	500	0.026	0.018	0.035	0.005	1	6655.84	6655.84	6655.84	***
	1000	0.024	0.016	0.034	0.004	2	6677.97	6644.61	6711.33	33.36
	1500	0.021	0.014	0.026	0.004	2	6614.26	6612.54	6615.97	1.72
	2000	0.022	0.018	0.027	0.003	3	6645.73	6627.55	6658.49	13.20
	2500	0.021	0.017	0.027	0.003	6	6658.29	6634.35	6685.84	16.75

Table 16: Solution results of AR(1) demand for production planning with pricing for $\theta = 0.02$

		Solution Risk				Profit for Feasible Solutions				
α	N	Avg	Min	Max	σ	#	Avg	Min	Max	σ
0	50	0.038	0.011	0.090	0.023	2	6425.48	6215.96	6635.01	209.52
	100	0.024	0.006	0.049	0.014	6	6453.05	6242.66	6632.97	154.77
	200	0.018	0.006	0.035	0.009	7	6365.02	6203.31	6576.07	123.50
	300	0.008	0.002	0.017	0.004	10	6261.79	5955.75	6414.22	127.06
0.02	1000	0.024	0.018	0.032	0.003	1	6602.79	6602.79	6602.79	***
	1500	0.022	0.018	0.026	0.002	2	6561.49	6546.36	6576.63	15.14
	2000	0.022	0.018	0.024	0.002	3	6567.83	6536.66	6610.64	31.30
	2500	0.022	0.018	0.028	0.003	4	6566.65	6529.12	6594.29	28.15

4.3.3 Upper bounds

We check the upper bounds for production planning with pricing options when $\alpha = \theta = 0.02$. The gaps are defined as the percent by which the upper bound is above the best feasible solution (i.e., the maximum profit among all feasible solutions in this case). We use $L = 1, \ldots, 4$ to generate bounds by the optimal objective values of the $M = 10$ SAA problems and the corresponding confidence levels given by (19) are 0.999, 0.989, 0.945, 0.828, respectively.

Table 17 shows the upper bounds for i.i.d. demand for production planning with pricing options when $\alpha = \theta = 0.02$. Table 18 shows the upper bounds for AR(1) demand for production planning with pricing options when $\alpha = \theta = 0.02$.

Combining the test results on the upper bounds and the results of feasible solutions in Section 4.3.2, we can obtain the range of the optimal profit. For example, in the i.i.d. demand case, solving $M = 10$ SAA instances with sample size $N = 1000$ yields a best feasible solution with profit 6711.33, as shown in Table 15; we also get an upper bound of 6740.62 with confidence 98.9% (shown in Table 17). This means that we have at least 98.9% confidence to say that the optimal profit is at most $(6740.62 - 6711.33)/6740.62 \approx 0.44\%$ greater than the best feasible solution 6740.62. We can make a similar analysis with other demand cases using corresponding tables.

From these results, we observe that as sample size N becomes larger, the upper bound becomes smaller and the gap becomes smaller at each confidence level. When the gap reaches zero, we can conclude that the best feasible solution is the optimal solution with the corresponding confidence level. For example, as we observe from Table 18, when the sample size $N = 1000$, we have confidence at least 94.5% to say that the optimal profit is at most 0.1% greater than 6609.35; when the sample size increases to $N = 2000$, we have confidence at least 94.5% to say that the feasible solution with profit 6610.64 is optimal. Thus, for a certain confidence level, we can make a better estimation of the optimal solution of an SAA problem as we increase the sample size N.

Table 17: Upper bounds of i.i.d. demand for production planning with pricing for $\alpha = \theta = 0.02$

	UB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
500	6741.58	6723.42	6695.46	6692.89	1.29%	1.02%	0.60%	0.56%
1000	6767.01	6740.62	6711.33	6688.30	0.83%	0.44%	0.00%	−0.34%
1500	6702.55	6693.99	6689.16	6668.40	1.31%	1.18%	1.11%	0.79%
2000	6687.78	6687.74	6670.77	6658.49	0.44%	0.44%	0.18%	0.00%
2500	6696.60	6696.56	6686.90	6685.84	0.16%	0.16%	0.02%	0.00%

Table 18: Upper bounds of AR(1) demand for production planning with pricing for $\alpha = \theta = 0.02$

	UB with confidence at least				Gap with confidence at least			
N	0.999	0.989	0.945	0.828	0.999	0.989	0.945	0.828
1000	6649.49	6621.39	6609.35	6606.17	0.71%	0.28%	0.10%	0.05%
1500	6659.58	6628.80	6623.04	6582.9	1.26%	0.79%	0.71%	0.10%
2000	6653.04	6636.38	6610.64	6605.67	0.64%	0.39%	0.00%	-0.08%
2500	6629.96	6613.79	6594.29	6593.37	0.54%	0.30%	0.00%	-0.01%

4.3.4 Sensitivity analysis

We conduct sensitivity analysis for production planning with pricing options. We focus on parameters a_t and β_t in the function $d_t(r_t) = a_t r_t + \beta_t + \tilde{\epsilon}_t$ for $t = 1, \ldots, T$. To better demonstrate our sensitivity results, we assume that the demand function is time invariant, i.e., $a_t = a$ and $\beta_t = \beta$ for all $t = 1, \ldots, T$. The noise term $\tilde{\epsilon}_t$ follows normal distribution with mean 0 and standard deviation 22 for all $t = 1, \ldots, T$. We use $\theta = 0.02$, $\alpha = 0$ and $N = 300$ since they can yield a feasible solution with high confidence level to the model, as shown in Table 15.

Sensitivity analysis on the slope a. We fix $\beta = 200$ and vary a in the set $\{-1, \ldots, -10\}$ in each period. Our test result is shown in Figures 1–3 in the Appendix.

As shown in Figure 1, the total profit increases as the absolute value of a decreases. Moreover, from Figure 2 and Figure 3, we observed that in each period, the optimal order quantity and the optimal price increase as the absolute value of a decreases.

The intuition is explained as follows. As the absolute value of a decreases, the demand will be less sensitive to the price; hence, the retailer has the motivation to increase the price while keeping the demand to be at least the same as before. As a result, the ordering quantities increase due to the increasing demands and the total profit also increases.

Sensitivity analysis on β. We fix $a = -5$ and vary β in the set $\{160, \ldots, 250\}$. Our test result is shown in Figures 4–6 in the appendix.

As shown in these three figures, the total profit, the optimal order quantity and the optimal price in each period all increase as the value of β increases.

As the value of β increases, the basic demand increases; hence, the retailer has the motivation to increase the price and still keep the demand at a higher level. The higher level demand leads to more order quantities. Consequently, the total profit increases.

5 Conclusions

In this paper, we propose two models of production planning problem under a probabilistic service-level guarantee (interpreted as stockout probabilities) over the entire planning horizon. The first model is an inventory management model while the second one also involves pricing decisions. We reformulate these two models as mixed-integer programs based on a finite set of discrete samples of the uncertainty, and solve them by using the SAA method. However, the resulting MILP models are computationally intractable since the SAA method requires a very large sample size; we computationally obtain feasible solutions and lower bounds on these models by adjusting the risk parameter, which gives us an efficient way to bound the optimal cost and the optimal profit. We conduct extensive computational tests under different service-level requirements and demand cases, so as to demonstrate the feasible solutions and lower bounds as well as to suggest reasonable sample sizes in practice.

An interesting direction for future research is to study the dynamic version of our model in which the price and inventory in each period can be changed dynamically at the beginning of each period. Such model solves dynamic pricing problems by using dynamic programming formulation rather than static MILP formulation. This is mainly due to the reason that MILP formulations for a dynamic problem require to grow a scenario tree to represent decisions based on the system states, which could be intractable to compute [see, e.g., 12, 13].

Acknowledgement. We sincerely thank the anonymous associate editor and three anonymous referees for their constructive comments and suggestions, which helped significantly improve both the content and the exposition of this paper. This research was funded in part by the United States National Science Foundation under grants CMMI-1433066 (Shen) and CMMI-1451078 (Shi). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

[1] Absi, N., S. Kedad-Sidhoum, S. Dauzère-Pérès. 2011. Uncapacitated lot-sizing problem with production time windows, early productions, backlogs and lost sales. *International Journal of Production Research* **49**(9) 2551–2566.

[2] ARI. 2015. Lead times, downtime and production delays. `http://www.arifleet.com/publications/lead_times_delays/#item1648`. Accessed August 2015.

[3] Bookbinder, J.H., J.-Y. Tan. 1988. Strategies for the probabilistic lot-sizing problem with service-level constraints. *Management Science* **34**(9) 1096–1108.

[4] CEMEX. 2015. Annual report 2014. http://www.cemex.com/CEMEX_AR2014/pdf/download/00_CX14OK.pdf. Accessed August 2015.

[5] Chen, F., J.-S. Song. 2001. Optimal policies for multiechelon inventory problems with markov-modulated demand. *Operations Research* **49**(2) 226–234.

[6] Chen, X., D. Simchi-Levi. 2012. *Pricing and inventory management, R. Philips and O. Ozer, eds.*. Oxford University Press, Oxford.

[7] Clark, A. J., H. Scarf. 1960. Optimal policies for a multi-echelon inventory problem. *Management science* **6**(4) 475–490.

[8] Fine, C.H., J.C. Lafrance, D. Hillebrand. 1996. The us automobile manufacturing industry. *Meeting the Challenge: US Industry Faces the 21st Century*. Citeseer, 35.

[9] Florian, M., J.K. Lenstra, Rinnooy K. 1980. Deterministic production planning: Algorithms and complexity. *Management science* **26**(7) 669–679.

[10] González-Ramírez, R.G., N.R. Smith, R.G. Askin. 2011. A heuristic approach for a multi-product capacitated lot-sizing problem with pricing. *International Journal of Production Research* **49**(4) 1173–1196.

[11] Gupta, A., C.D. Maranas. 2003. Managing demand uncertainty in supply chain planning. *Computers & Chemical Engineering* **27**(8) 1219–1227.

[12] Huang, K., S. Ahmed. 2005. The value of multi-stage stochastic programming in capacity planning under uncertainty. *Stochastic Programming E-Print Series* **15**.

[13] Huang, K., S. Ahmed. 2008. On a multi-stage stochastic programming model for inventory planning. *INFOR: Information Systems and Operational Research* **46**(3) 155–163. doi:10.3138/infor.46.3.155.

[14] Kazaz, B. 2004. Production planning under yield and demand uncertainty with yield-dependent cost and price. *Manufacturing & Service Operations Management* **6**(3) 209–224.

[15] Kleywegt, A.J., A. Shapiro, T. Homem-de Mello. 2002. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* **12**(2) 479–502.

[16] Luedtke, J., S. Ahmed. 2008. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization* **19**(2) 674–699.

[17] Mancilla, C., R. Storer. 2012. A sample average approximation approach to stochastic appointment sequencing and scheduling. *IIE Transactions* **44**(8) 655–670.

[18] McCormick, G.P. 1976. Computability of global solutions to factorable nonconvex programs: Part i–convex underestimating problems. *Mathematical programming* **10**(1) 147–175.

[19] Mills, Edwin S. 1959. Uncertainty and price theory. *The Quarterly Journal of Economics* 116–130.

[20] Mills, T.C. 1991. *Time Series Techniques for Economists*. Cambridge University Press.

[21] Mula, J., R. Poler, J.P. Garcia-Sabater, F.C. Lario. 2006. Models for production planning under uncertainty: A review. *International journal of production economics* **103**(1) 271–285.

[22] Pagnoncelli, B.K., S. Ahmed, A. Shapiro. 2009. Sample average approximation method for chance constrained programming: theory and applications. *Journal of optimization theory and applications* **142**(2) 399–416.

[23] Pochet, Y., L.A. Wolsey. 1988. Lot-size models with backlogging: Strong reformulations and cutting planes. *Mathematical Programming* **40**(1) 317–335.

[24] Verweij, B., S. Ahmed, A.J. Kleywegt, G. Nemhauser, A. Shapiro. 2003. The sample average approximation method applied to stochastic routing problems: a computational study. *Computational Optimization and Applications* **24**(2-3) 289–333.

[25] Zangwill, W. I. 1969. A backlogging model and a multi-echelon model of a dynamic economic lot size production system – a network approach. *Management Science* **15**(9) 506–527.

[26] Zipkin, P. 2000. *Foundations of Inventory Management*, vol. 2. McGraw-Hill, New York.

Appendix

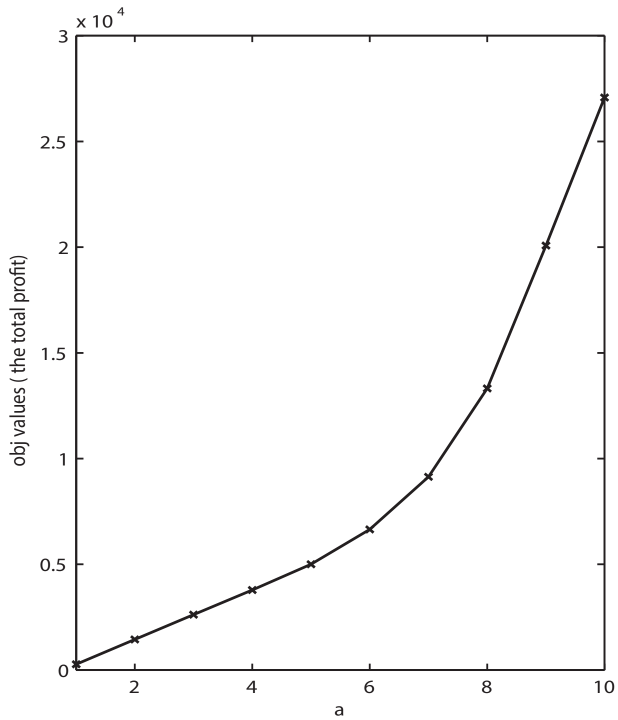

Figure 1: Sensitivity analysis on the slope a: objective values

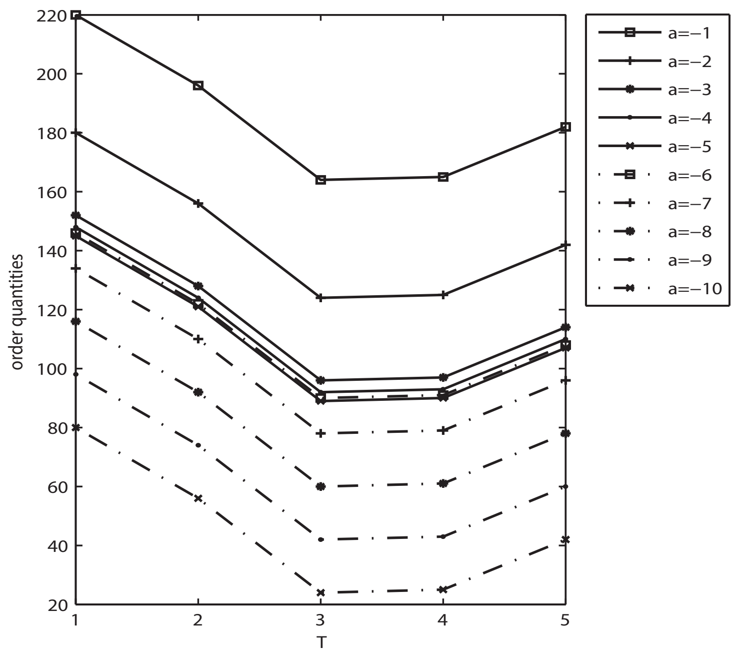

Figure 2: Sensitivity analysis on the slope a: order quantities

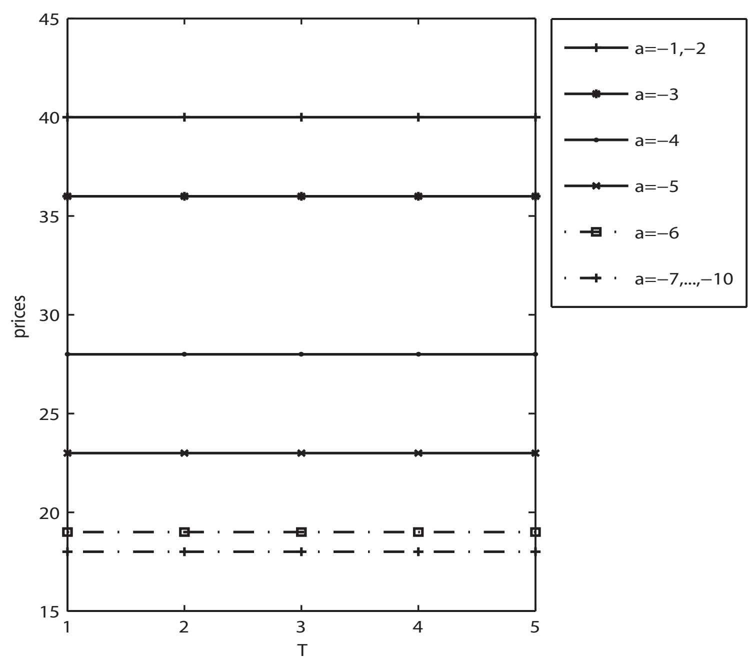

Figure 3: Sensitivity analysis on the slope a: prices

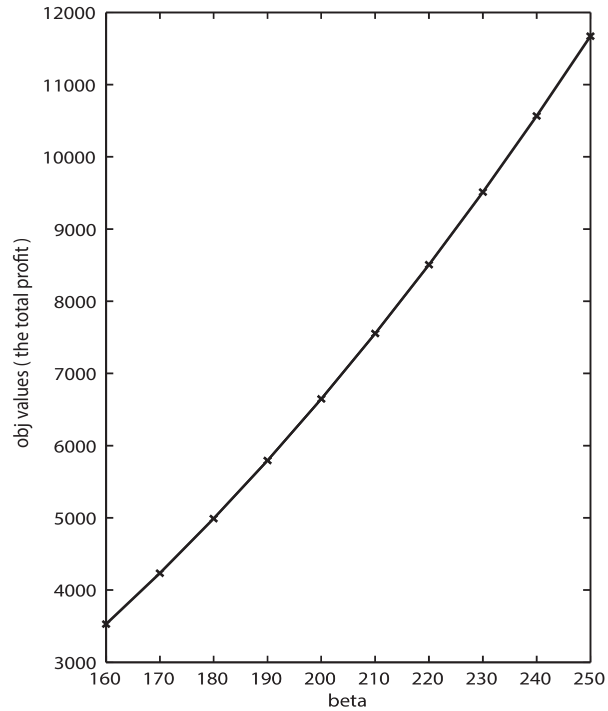

Figure 4: Sensitivity analysis on β: objective values

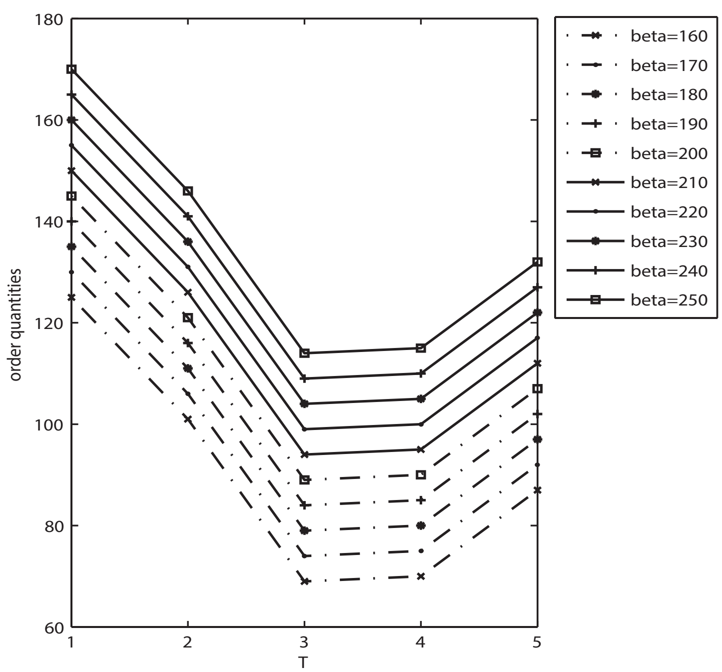

Figure 5: Sensitivity analysis on β: order quantities

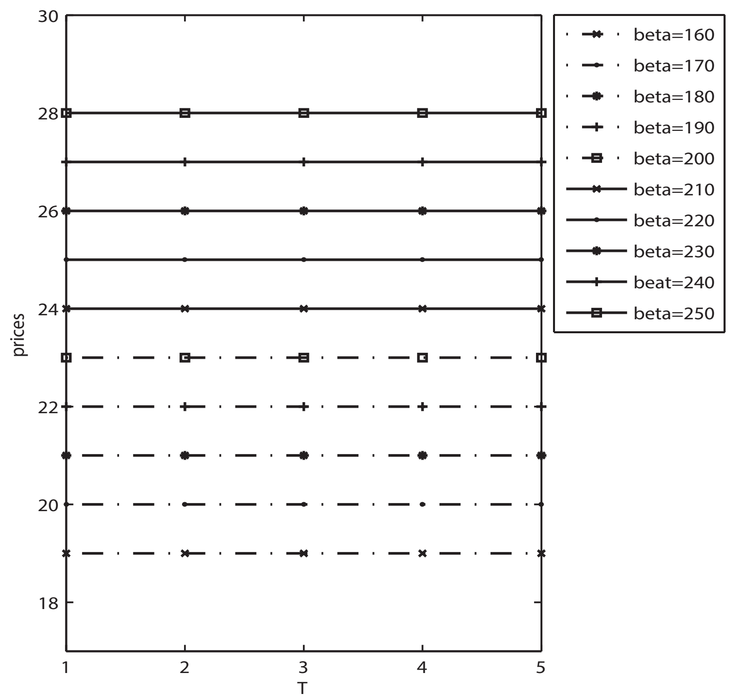

Figure 6: Sensitivity analysis on β: prices