On minimizing downside risk in make-to-stock, risk-averse firms

Bo Li1,2 | Antonio Arreola-Risa3

1Line-haul Road Transportation Lab, Ningbo China Institute for Supply Chain Innovation, Ningbo, Zhejiang, China
2Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, Massachusetts,
3Information and Operations Management Department, Mays Business School, Texas A&M University, College Station, Texas,

Correspondence
Bo Li, Ningbo China Institute for Supply Chain Innovation, No. 169 South Qixing Road, Meishan, Ningbo, Zhejiang 315832, China.
Email: libo@mit.edu

Funding information
Natural Science Foundation of Ningbo, Grant/Award Numbers: 2018A610120, 2019A610047. Zhejiang Qianjiang Talent Program, Grant/Award Number: QJC1803003. NISCI-TDH Line-haul Road Transportation Lab.

Abstract
Consider a single-product, make-to-stock, risk-averse firm with stochastic demand and manufacturing capacity. Temporarily unfulfilled demands are back-ordered. Total cost is equal to the sum of inventory-holding and back-ordering costs over a continuous-time, finite planning horizon. In this article, we study the problem of finding the base-stock level that minimizes total cost conditional value-at-risk (CVaR), or total cost CVaR for short. We demonstrate that total cost CVaR is both convex in the base-stock level and increasing in risk aversion. We derive a closed-form approximation of total cost which is asymptotically exact in the planning horizon length. The approximation leads to a relatively simple determination of optimal base-stock levels which are reasonably accurate for practical applications. We make observations and identify findings regarding the impact on optimal base-stock levels of changes in risk aversion, manufacturing capacity, inventory-holding and back-ordering costs, and planning horizon length. We provide a detailed description of how our research results were applied in a real-world, supply-chain design project.

KEYWORDS
base-stock policy, conditional value-at-risk, downside risk, make-to-stock, stochastic demand and capacity

1 | INTRODUCTION

The likelihood of a random cost exceeding a threshold value is a basic definition of downside risk in financial literature. Rockafellar and Uryasev (2002) propose a measure of downside risk called conditional value-at-risk (CVaR) which has quickly gained popularity in the investment and insurance communities. For a given cost probability distribution, CVaR is defined as the expected cost in the top \((1 - \beta)\) percentile of the distribution, where \(\beta \in (0, 1)\). The higher the value of \(\beta\), the higher the firm’s risk aversion (Chen, Xu, & Zhang, 2009). In this article, we focus on CVaR for two major reasons. First, most firms are risk averse (Amihud & Lev, 1981; Dave, Eckel, Johnson, & Rojas, 2010; Schweitzer & Cachon, 2000) and, consequently, they are sensitive to the downside risk measured by CVaR. Second, CVaR is both a coherent risk measure (Rockafellar & Uryasev, 2002) and generally leads to analytical tractability.

Consider a single-product, make-to-stock firm. The product is demanded one unit at a time in a stochastic and stationary fashion. The product is also manufactured one unit at a time, unit manufacturing time is stochastic, and in consequence, manufacturing capacity (i.e., the number of units that can be produced in a time period) is stochastic as well. To buffer the random interaction of demand and manufacturing, the firm keeps an inventory of the product, which is managed by a base-stock policy with parameter \(S\). See DeCroix and Arreola-Risa (1998) for the optimality of a base-stock policy in a setting similar to ours. The base-stock policy operates as follows: Starting with an on-hand inventory of \(S\) units, each arriving demand for one unit triggers a production order for one unit, and consequently, the company stops production when on-hand inventory reaches \(S\). Demands which arrive when the inventory is temporarily out of stock are back-ordered, and they will be fulfilled on a first-come, first served basis as production orders are completed by the manufacturing system.

Each unit on hand will incur an inventory-holding cost per period, and each unfilled demand will incur a back-ordering cost per period. Let us define the random variable total cost as the sum of inventory-holding and back-ordering costs over a continuous-time, finite planning horizon. As a consequence...
of demand and manufacturing capacity being stochastic, total cost is a random variable. We will shorten CVaR of total cost to total cost CVaR. In this paper, we address the problem of finding the base-stock level \( S^* \) which minimizes total cost CVaR. This problem will be referred to as the problem under study, and \( S^* \) will be called the optimal base-stock level. Our interest in this problem originates from the research we conducted in a supply-chain design project at a Fortune 20 energy company. For confidentiality, the company will be identified as Company Delta.

We make four contributions with this article. First, we demonstrate that total cost CVaR is convex in \( S \), increasing in \( \beta \), and jointly convex in per-period unit inventory-holding and back-ordering costs. Second, we derive closed-form approximations of expected total cost, variance of total cost, and total cost CVaR, which are all asymptotically exact in planning horizon length. These approximations yielded optimal or near-optimal base-stock levels in a wide variety of test problems.

The third contribution of the paper is the identification of observations and the development of findings regarding the impact on optimal base-stock levels of changes in risk aversion, manufacturing capacity, per-period unit inventory-holding cost, per-period unit back-ordering cost, and planning horizon length.

The fourth contribution of the paper is to provide a detailed description of how our research results were applied in a real-world supply-chain design project. Due to our confidentiality agreement with Company Delta, the information in the description is realistic but not real. However, the applications provided should be valuable learning experiences for researchers and practitioners alike.

The remainder of the paper is divided into seven sections. Section 2 positions our paper in the related published literature. Section 3 presents a model along with a formal definition of the problem under study. In Section 4 we establish that total cost CVaR is convex in \( S \), increasing in \( \beta \), and jointly convex in per-period unit inventory-holding and back-ordering costs. Additionally, in Section 4 we derive closed-form approximations of expected total cost and variance of total cost, which are asymptotically exact in planning horizon length. In Section 5, we derive a closed-form approximation of total cost CVaR, which is asymptotically exact in planning horizon length, and assess the approximation’s accuracy in a wide variety of test problems. Section 6 presents a sensitivity analysis which leads to observations and findings regarding the impact on the optimal base-stock level of changes in risk aversion, manufacturing capacity, per-period unit inventory-holding cost, per-period unit back-ordering cost, and planning horizon length. Section 7 illustrates how some of the research results in this paper were applied in Company Delta’s supply-chain design project. The final section summarizes our research results and lists future research ideas.

2 LITERATURE REVIEW

CVaR is a fairly new downside risk measure, also known as expected shortfall, average value at risk, and expected tail loss. CVaR possesses many desirable theoretical properties such as monotonicity, sub-additivity, positive homogeneity, and translation invariance (Rockafellar & Uryasev, 2002). These properties make CVaR an excellent and cohesive representation of subjective risk aversion. Moreover, CVaR is consistent with observed biological foundations of loss aversion (Tom, Fox, Trepel, & Poldrack, 2007).

Consideration of downside risk in general and CVaR in particular is not new in the operations and supply-chain management literature (Rao & Goldsby, 2009). For example, several papers have adopted CVaR as the objective function in highly uncertain environments such as those found in dual-sourcing strategies (Tomlin & Wang, 2005), channel coordination with multiple risk-averse suppliers (Chen, Shum, & Simchi-Levi, 2014), and resource allocation (Wagner & Radovilsky, 2012).

Our paper falls at the intersection of two major and independent streams of research in operations management. In the first research stream, downside risk is minimized in risk-averse firms which either purchase the product from suppliers, produce the product in a manufacturing system with infinite capacity, and/or make a single inventory-production decision at the beginning of a one-period planning horizon. For example, Gotth and Takano (2007) consider a one-period newsvendor model with infinite capacity and find the order quantity minimizing CVaR. Chen et al. (2009) analyze the pricing and ordering decisions in one-period uncapacitated newsvendor models with additive or multiplicative demands and a CVaR criterion. Wu, Zhu, and Teunter (2013) study one-period risk-averse newsvendor models with random capacity under both value-at-risk and CVaR criteria. One can also find attempts to extend the application of CVaR from one-period newsvendor models to multi-period inventory models (Borgonovo & Pecciati, 2009; Zhang, Xu, & Wu, 2009). In related research, Li and Arreola-Risa (2017) and Serrano, Oliva, and Kraiselburd (2017) examine inventory management to maximize firm market valuation.

In the second research stream, either the make-to-stock system is capacitated or multiple inventory-production decisions are made over a finite (or infinite) planning horizon, or both; however, the firm is risk-neutral, and instead of minimizing downside risk, the objective is to minimize an expected value. For example, the seminal Kaplan (1970) paper deals with the characterization of the optimal inventory policy for a multistage, uncapacitated inventory system in which the objective function is the present value of the expected holding and back-ordering costs. Years later, Karmarkar (1987) proposes a capacitated make-to-stock model of a risk-neutral firm in which the supply source is a single-server queue which is affected by the order quantity. Independently, Zipkin (1986) considers an analogous situation to Karmarkar’s,
but now Zipkin models the supply source as a queuing network. In later work, Lee and Zipkin (1992, 1995) examine supply chain settings for risk-neutral firms in which the supply source is modeled as serial queues and networks of queues, respectively. More recently, Arreola-Risa (1996) analyzes an infinite-horizon supply chain model of a risk-neutral firm in which the production system is shared by multiple products, and lead time depends on both product demand and manufacturing capacity.

Our paper is related to the first research stream, because the paper considers make-to-stock firms that produce a product in a manufacturing system with finite capacity. Moreover, our paper is related to the second research stream, because the paper considers downside risk for a risk-averse firm. Because our paper combines downside risk with a continuous-time, finite planning horizon in a capacitated make-to-stock system, it is now possible to study how manufacturing capacity, planning horizon length, and a firm’s risk aversion affect the optimal base-stock level. Equally important, our research results can be used to capture the inherent trade-offs among risk aversion, inventory, manufacturing capacity, and planning horizon length when the objective is to minimize downside risk.

To close this section, we will cite Carneiro, Ribas, and Hamacher (2010). That publication, like the problem under study, was motivated by research at a real-world energy company which utilizes the CVaR downside risk measure to manage supply chain risk. However, the problem and the model in Carneiro et al. (2010) are different from our study in two fundamental ways. In their problem, the goal is to find an optimal portfolio of investments, while our problem is to find an optimal base-stock level. At the same time, their model is a two-stage stochastic model with fixed recourse, while ours is a Markov reward process model.

### 3 | MODEL AND PROBLEM DEFINITION

Because the product is demanded one unit at a time in a stochastic and stationary fashion, following common practice (see, eg, Zipkin, 1986; Arreola-Risa, 1996), we model demand by a Poisson process. Let $D(t)$ denote demand by time $t$, $p_X(\cdot)$ denote the probability mass function of random variable $X$, and $\mathbb{E}[\cdot]$ denote the expected value operator. Thus, when $D(t)$ is a Poisson process, $p_{D(0)}(x) = \frac{e^{-\mu(t)} \mu^x}{x!}$, $x = 0, 1, \ldots$ (1) and $\mathbb{E}[D(t)] = \lambda t$, (2) where the parameter $\lambda$ represents the average demand per unit time, or equivalently, the average demand rate.

Unit manufacturing times are independent and identically distributed (IID) continuous random variables. Thus, manufacturing capacity, measured as the number of units that can be produced in a time period, is stochastic as well. Let $M$ denote unit manufacturing time. Again, following common practice (see, eg, Zipkin, 1986; Arreola-Risa, 1996), we model $M$ by an exponential random variable. Let $f_X(\cdot)$ denote the probability density function of random variable $X$. Then, when $M$ is an exponential random variable, $f_M(t) = \mu e^{-\mu t}, \quad t > 0$ (3) and $\mathbb{E}[M] = \frac{1}{\mu}$, (4) where the parameter $\mu$ represents average manufacturing rate, or equivalently, $1/\mu$ represents average unit manufacturing time.

See the notations in Table 1. The decision variable $S$ is included in the notation to emphasize dependence on $S$. To justify the use of a time-invariant base-stock level in a finite planning horizon setting, we will invoke the well-known result in Veinott (1965) and assume that any units on-hand or back-ordered at the end of the planning horizon can be salvaged or filled at the production cost. Moreover, following related research with long planning horizons, we assume that the make-to-stock system starts in an initial state which follows its steady state distribution.

| Table 1 Notations |  |
|-------------------|---|
| Notation | Definition |
| $h$ | Unit inventory-holding cost per period |
| $b$ | Unit back-ordering cost per period |
| $I(S, t)$ | Instantaneous cost rate at time $t$ |
| $T$ | Finite planning horizon length |
| $K(S)$ | Total cost (sum of the inventory-holding and back-ordering costs in $[0, T]$) |
| $f_K(S, \cdot)$ | Probability density function of $K(S)$ |
| $F_K(S, \cdot)$ | Cumulative probability distribution of $K(S)$ |

Define now $\eta_\beta(S)$ as the lowest amount such that with probability $\beta$ the total cost will be no higher than $\eta_\beta(S)$. Then, $\eta_\beta(S) = \min\{\eta \in \mathbb{R} : F_K(S, \eta) \geq \beta\}$. (7)

Inclusion of $\beta$ and $S$ in the notation emphasizes dependence on the parameter $\beta$ and the base-stock level $S$.

Recall that total cost CVaR equals the expected total cost in the top $(1 - \beta)$ percentile of a total cost probability distribution. Let $\zeta_\beta(S)$ denote total cost CVaR as a function of $\beta$ and $S$. A moment’s reflection then reveals that

$$\zeta_\beta(S) = (1 - \beta)^{-1} \int_{y \geq \eta_\beta(S)} y f_K(S, y) dy.$$ (8)

To aid the reader’s intuition, in Figure 1 we present a representative example of a probability distribution of total cost and a total cost CVaR for selected values of $\beta$ and $S$. This example comes from the numerical experiments in Section 5.
Recall that $S^*$ is the minimizer of $\xi_\beta(S)$. We can now formally and succinctly define the problem under study as follows: Given $\beta, h, b, \rho$, find $S^*$ when $D(t)$ is a Poisson process with parameter $\lambda$, and $M$ is an exponential random variable with parameter $\mu$.

4 | TOTAL COST

It is clear from Equation (8) that, in our search for $S^*$, we first need to determine $f_K(S, \cdot)$. Recall that $I(S, t)$ is the instantaneous cost rate at time $t$. Letting $p_{I(S,t)}(\cdot)$ denote the probability mass function of $I(S, t)$, it is easy to see that in principle $f_K(S, \cdot)$ can be obtained by taking convolutions of $p_{I(S,t)}(\cdot)$ over $[0, T]$. The functional form of $p_{I(S,t)}(\cdot)$ is established in the following lemma. Define $\rho \equiv \lambda/\mu$ as the capacity utilization of the make-to-stock system. All proofs are deferred to the Appendices (Supporting information).

**Lemma 1**

$$p_{I(S,t)}(y) = \begin{cases} 0, & \text{otherwise.} \\ (1 - \rho)\rho^{x+y/h}, & \text{if } y/h \in \mathbb{N}, \ y/b \not\in \mathbb{N} \text{ and } y/h \in [0, S]; \\ (1 - \rho)\rho^{x+y/b} + \rho^{x+y/h}, & \text{if } y/b \in \mathbb{N}, \ y/h \in \mathbb{N}, \text{and } y/h \in [0, S]; \\ (1 - \rho)\rho^{x+y/b}, & \text{if } y/b \not\in \mathbb{N} \text{ and } y/h \not\in \mathbb{N}; \end{cases}$$

Note that the functional form of $p_{I(S,t)}(\cdot)$ is independent of $t$. This is the case because, as explained in the proof of Lemma 1, in steady state the number of outstanding production orders is independent of $t$, which causes $I(S, t)$ to be independent of $t$ as well.

Unfortunately, integrating $p_{I(S,t)}(\cdot)$ over $[0, T]$ to obtain $f_K(S, \cdot)$ is intractable due to the auto-correlation in consecutive instantaneous costs emanating from the auto-correlation in the number of outstanding production orders. Fortunately, building on Lemma 1, we are able to characterize $\xi_\beta(S)$ in Theorem 1.

**Theorem 1** Regarding total cost CVaR, we have that

(a) $\xi_\beta(S)$ is convex in $S$

(b) $\xi_\beta(S)$ is increasing in $\beta$

(c) $\xi_\beta(S)$ is jointly convex in $h$ and $b$.

Theorem 1(a) guarantees the existence of a global minimum total cost CVaR. At the same time, Theorem 1(b) establishes that a higher risk aversion will lead to a higher minimum total cost CVaR. Lastly, Theorem 1(c) establishes that there is a nonlinear impact of the cost parameters $h$ and $b$ on total cost CVaR.

The convexity properties established in Theorem 1 guarantee the success of using simulation and numerical analysis to minimize total cost CVaR. However, given the well-known computation inefficiencies of simulation, as well as its limitations to produce theoretical and managerial observations and findings, we will now focus our attention on a two-moment analytical approximation of total cost CVaR. To this end, in Sections 4.1 and 4.2 we will respectively derive closed-form approximations of expected total cost and of variance of total cost. The approximations are asymptotically exact in $T$. The strategy to obtain the approximations is to exploit the fact that the evolution over time of total cost corresponds to the evolution of an aperiodic and recurrent Markov chain with an infinite number of states and a unique stationary distribution. In particular, we will use Markov reward process theory (Sladký & van Dijk, 2005; van Dijk & Sladký, 2006). Although Sladký and van Dijk (2005) and van Dijk and Sladký (2006) consider finite-state Markov chains, their technique is also applicable to infinite-state, aperiodic Markov chains with a unique stationary distribution, which is precisely the case for the evolution of $K(S)$.

4.1 | Expected total cost

Define the number of outstanding production orders at the manufacturing facility as the state of the make-to-stock system being optimized. Let $\lambda(t)$ be the state of the make-to-stock system at time $t$, $K(S, t)$ be the total cost up to time $t$, and $R(t) = \mathbb{E}[K(S, t) | X(0) = i]$ be the expected total cost up to time $t$, when starting from state $i$. According to van Dijk and Sladký (2006), $R(t)$ will grow linearly with time at a constant rate, say $g$, plus a constant, say $w_i$, and plus a term $\epsilon(t)$, where the term $\epsilon(t)$ will converge exponentially quickly to 0 as $t \to \infty$. In other words, $R(t) = g \cdot t + w_i + \epsilon(t)$ and $R_i(t) \to g \cdot t + w_i$ as $t \to \infty$. Interested readers may find expressions of $w_i$ in Remark 1 of the Appendices.
We know from Equation (5) that expected total cost is $\mathbb{E}[K(S)]$. In the Proposition 1 we establish a closed-form approximation of $\mathbb{E}[K(S)]$ which is asymptotically exact in $T$.

**Proposition 1** As $T \to \infty$, $\frac{\mathbb{E}[K(S)]}{T} \to g(S)$, where

$$g(S) = h \left( S - \frac{\rho}{1 - \rho} \right) + (h + b) \frac{\rho^{S+1}}{1 - \rho}. \quad (10)$$

It can be trivially argued that, all else being equal, $g(S)$ is increasing in $h$, $b$, and $\rho$.

A risk-averse firm with $\beta = 0$ is equivalent to a risk-neutral firm, because substituting $\beta = 0$ in Equations (7) and (8) yields $\zeta_\beta(S) = \mathbb{E}[K(S)]$. In the proposition that follows, we establish the optimal base-stock level of a risk-neutral firm which we denote by $S^*_g$.

**Proposition 2** The function $g(S)$ in Proposition 1 is convex in $S$, and its minimizer is given by

$$S^*_g = \left[ \frac{\ln(h) - \ln(h + b)}{\ln(\rho)} \right]. \quad (11)$$

Proposition 2 is valuable not only on its own right, but also because it can be trivially argued that all else being equal, $S^*_g$ is increasing in both $b/h$ and $\rho$. Later in the paper, Equation (11) will play a crucial role in characterizing the behavior of $S^*$ as the values of $h$, $b$, $\lambda$, and $\mu$ change. Figure 2 presents several examples which illustrate the convexity of $g(S)$ in $S$ and that $S^*_g$ is increasing in the ratios $b/h$ and $\rho \equiv \lambda/\mu$.

### 4.2 Variance of total cost

Let $\mathbb{V}[-]$ be the variance operator and $\mathbb{V}[K(S)]$ denote the variance of total cost. Using once more Markov reward process theory, we will derive a closed-form approximation of $\mathbb{V}[K(S)]$ which is asymptotically exact in $T$. As a preliminary step in our effort to determine $\mathbb{V}[K(S)]$, we first establish a lemma which contains analytical expressions of the differences $w_i - w_{i-1}$.

**Lemma 2**

(i) If $1 \leq i \leq S + 1$, then

$$w_i - w_{i-1} = \frac{1}{\lambda} \left[ h \left( \frac{-\rho}{1 - \rho} \right) + (h + b) \frac{\rho^{S+2-i} - \rho^{S+2}}{(1 - \rho)^2} \right].$$

(ii) If $i \geq S + 2$, then

$$w_i - w_{i-1} = \frac{1}{\lambda} \left\{ b(S + 1) \left( \frac{-\rho}{1 - \rho} \right) + h \frac{\rho - \rho^{S+2}}{(1 - \rho)^2} \right. \\
- \left. \frac{b((i - S) - 1)\rho^2 - (i - S)\rho + \rho^{S+2}}{(1 - \rho^2)} \right\}. \quad (12)$$

Although Equation (12) is a transcendental function difficult to analyze, in the next corollary of Proposition 3 we derive six properties of the equation. For easy reference, the properties are called Properties 1 to 6.

**Corollary 1** Regarding $\gamma(S)$, we have that

1. $\frac{\partial^2}{\partial \rho^2} \gamma(S) \bigg|_{S=0} = 0$ and $\frac{\partial^2}{\partial \rho^2} \gamma(S) \bigg|_{S \geq 1} > 0$.
2. $\frac{\partial^2}{\partial \rho^2} \gamma(S) > 0$.
3. $\frac{\partial^2}{\partial h^2} \gamma(S) < 0$.
The limit of $\gamma(S)$ when $S \to +\infty$ is

$$\lim_{S \to +\infty} \gamma(S) = 2h^2 \mu^{-1}(1 - \rho)^{-1} \rho(1 + \rho).$$  

The value of $\gamma(S)$ when $S = 0$ is

$$\gamma(S)|_{S=0} = 2b^2 \mu^{-1}(1 - \rho)^{-1} \rho(1 + \rho).$$  

All else being equal, including $\rho$, $\gamma(S)$ decreases in $\mu$, and at the same time, $g(S)$ is independent of $\mu$.

Property 1 means that, except for the special case $S = 0$, a higher $h$ brings more than a linear increment in $\gamma(S)$, since $\gamma(S)$ is strictly convex in $h$ when $S \geq 1$ and always convex in $h$. Property 2 means that a higher $b$ brings more than a linear increment in $\gamma(S)$ since $\gamma(S)$ is strictly convex in $b$. Property 3 indicates a substitution effect between $h$ and $b$ in their contribution to $\gamma(S)$. In other words, the contribution of $b$ will be subduced as $h$ rises. At the same time, the contribution of $h$ will be subduced as $b$ rises. An implication of Property 3 is that the impact of $h$ and $b$ on the variance of total cost depends on each other in a nonreinforcing way. Note that this dependency is absent in the expected total cost, and we will see later, this dependency will be present in total cost CVaR.

Regarding Properties 4 and 5, since $S$ serves as a buffer between stochastic demand and stochastic capacity, then when $S$ takes its minimum value of zero and, hence, there is no buffer, the on-hand inventory will always be zero. Therefore, the variance of total cost will exclusively come from back-orders and their associated back-ordering cost, which is confirmed by Equation (14). On the other hand, as $S$ grows without bound, average back-orders will converge to zero. Consequently, the variance of total cost will exclusively come from units on hand and their associated holding cost, which is confirmed by Equation (13). We find it fascinating that in these two extreme and opposite values of $S$, the functional form of $\gamma(S)$ is identical, with the only difference being the terms $h^2$ and $b^2$. Because of the identical functional forms, the ratio of their respective variances of total cost is a constant,

$$\lim_{S \to +\infty} \frac{\gamma(S)}{\gamma(S)|_{S=0}} = \frac{h^2}{b^2}.$$  

Property 6 states that all else being equal, if $\mu_A < \mu_B$, Firms A and B will have the same expected total cost, but Firm B will have a lower variance of total cost. In other words, all else being equal, the variance of total cost in firms with faster manufacturing rates will be lower than the variance of total cost in firms with slower manufacturing rates.

5 TOTAL COST CVaR

In Theorem 2, we will use Propositions 1 and 3 to derive a closed-form approximation of total cost CVaR, which is asymptotically exact in $T$. Let $\phi(\cdot)$ and $\Phi(\cdot)$ respectively denote the probability density function and the cumulative density function of the standard normal distribution.

**Theorem 2** As $T \to \infty$,

$$\zeta_\beta(S) \to g(S) + \kappa_\beta \sqrt{\gamma(S)|T},$$

where

$$\kappa_\beta = \frac{\phi(\Phi^{-1}(\beta))}{1 - \beta}$$

and $g(S)$ and $\gamma(S)$ are respectively given in Equations (10) and (12).

Note in Equation (16) that the notation $\kappa_\beta$ is used to emphasize that the constant $\kappa$ is a function of $\beta$. Also note in Theorem 2 that in the limit when $T \to \infty$, $\zeta_\beta(S) = g(S) + \o(T) \to g(S)T$. This is a natural consequence of as $T \to \infty$, $K(S)T \to E[K(S)]T \to g(S)T$. Moreover, since the gap between $E[K(S)]$ and $\zeta_\beta(S)$ is $\kappa_\beta \sqrt{\gamma(S)|T}$ according to Theorem 2, the gap between $E[K(S)]$ and $\zeta_\beta(S)$ as a percentage of $E[K(S)]$ diminishes as $T \to \infty$. Therefore, we arrive at the following insight:

**Insight 1.** The base-stock levels minimizing $\zeta_\beta(S)$ and $E[K(S)]$ converge as $T \to \infty$.

Guided by Theorem 2, we propose to find the optimal base-stock level $S^*$ by replacing Equation (8) with the much simpler equation

$$\zeta_\beta(S) \approx \tilde{\zeta}_\beta(S) \equiv g(S)T + \kappa_\beta \sqrt{\gamma(S)|T}$$

in which $\tilde{\Delta}$ denotes an approximation of $\Delta$. Moreover, the following corollary of Theorem 2 sheds light on the impact of a firm’s risk aversion on total cost CVaR.

**Corollary 2** $\kappa_\beta$ increases in $\beta$ and thus in risk aversion.

We know that the higher the $\beta$, the higher a firm’s risk aversion. Corollary 2 establishes that in the approximation $\tilde{\zeta}_\beta(S)$, because $g(S)$ and $\gamma(S)$ are independent of $\beta$, increases in total cost CVaR due to increases in $\beta$ exclusively emanate from $\kappa_\beta$. We think that although $\tilde{\zeta}_\beta(S)$ is an approximation, being able to quantify the magnitude of the $\tilde{\zeta}_\beta(S)$ increases via $\kappa_\beta$ is of great value.

5.1 Accuracy of the total cost CVaR approximation in finding optimal base-stock levels

To better understand the accuracy of Equation (17) as $T$ increases, we conducted a numerical and simulation
experiment consisting of 54 test problems. The test problems were created by setting $\mu = 10$, $h = 1$, and then taking all combinations of the following parameter values:

- $\beta = 0.90$ and 0.95
- $\rho = 0.5$, 0.7, and 0.9
- $b/h = 1$, 5, and 25
- $T = 10$, 100, and 1000.

The two $\beta$ values are its most common values (see, eg, Rockafellar & Uryasev, 2002). The three $\rho$ values represent low, medium, and high values (see, eg, U.S. Board of Governors of the Federal Reserve System, 2019). The three $b/h$ values represent low, medium, and high values (see, eg, Arreola-Risa, 1996). The three values of $T$ represent short, medium, and long planning horizons.

Note that by setting $\mu = 10$, we can focus on capacity utilization $\rho$, which is more general than specific values of $\lambda$ and $\mu$. Similarly, by setting $h = 1$, we can focus on the $b/h$ ratio, which is more general than specific values of $h$ and $b$. The 54 test problems are presented in Table 2, in which the convergence discrepancy in percentage is explained in the next paragraph.

For each test problem, we simulated the make-to-stock system 100 000 times to obtain 100 000 total cost observations. In each simulation, we randomly selected the starting number of outstanding production orders from its steady-state geometric distribution with parameter $\rho$ (see Appendix 1). We then utilized the 100 000 total cost observations to construct a total cost probability distribution. Using the constructed total cost probability distribution, we performed an exhaustive search on the base-stock level $S$ in order to find $S^*$ and its associated $\zeta_\beta(S^*)$. Next, we used Equation (17) to analytically determine the minimizer of $\hat{\zeta}_\beta(S)$, which from now on will be denoted by $\hat{S}^*$. Lastly, using the $\hat{S}^*$ and the constructed total cost probability distribution, we calculated $\zeta_\beta(\hat{S}^*)$ and then computed the following convergence discrepancy as a percentage:

$$\frac{\zeta_\beta(\hat{S}^*) - \zeta_\beta(S^*)}{\zeta_\beta(S^*)}.$$  

Obviously, $\zeta_\beta(\hat{S}^*) \geq \zeta_\beta(S^*)$. 

**FIGURE 3**  Examples of the long-run per-period variance of total cost as a function of the base-stock level [Colour figure can be viewed at wileyonlinelibrary.com]
TABLE 2  The 54 test problems and their convergence discrepancies in percentage

| $\beta$ | $\rho$ | Test problem | $b/h$ | $T = 10$ | $T = 100$ | $T = 1000$ |
|---------|--------|--------------|-------|---------|---------|---------|
| 0.90    | 0.5    | 1-3          | 1     | 0.00    | 0.00    | 0.00    |
|         |        | 4-6          | 5     | 5.53    | 0.73    | 0.00    |
|         |        | 7-9          | 25    | 0.00    | 1.27    | 0.00    |
| 0.7     | 0.7    | 10-12        | 1     | 1.60    | 0.00    | 0.00    |
|         |        | 13-15        | 5     | 0.00    | 0.19    | 0.00    |
|         |        | 16-18        | 25    | 12.54   | 0.29    | 0.00    |
| 0.9     | 0.9    | 19-21        | 1     | 1.70    | 2.57    | 0.00    |
|         |        | 22-24        | 5     | 5.71    | 0.00    | 0.52    |
|         |        | 25-27        | 25    | 23.15   | 4.90    | 2.62    |
| 0.95    | 0.5    | 28-30        | 1     | 0.00    | 0.00    | 0.00    |
|         |        | 31-33        | 5     | 1.08    | 0.00    | 0.00    |
|         |        | 34-36        | 25    | 2.70    | 3.50    | 0.00    |
| 0.7     | 0.7    | 37-39        | 1     | 0.00    | 0.00    | 0.00    |
|         |        | 40-42        | 5     | 0.00    | 0.00    | 0.00    |
|         |        | 43-45        | 25    | 4.10    | 1.52    | 0.40    |
| 0.9     | 0.9    | 46-48        | 1     | 0.95    | 5.29    | 0.00    |
|         |        | 49-51        | 5     | 0.82    | 3.13    | 0.83    |
|         |        | 52-54        | 25    | 15.92   | 1.30    | 6.04    |

For example, when $\beta = 0.95$, $\rho = 0.7$, and $b/h = 5$, which corresponds to test problems 40, 41, and 42, we respectively obtained $S^* = \hat{S}^* = 10$, $S^* = \hat{S}^* = 7$, and $S^* = \hat{S}^* = 5$. Consequently, in test problems 40 to 42, Equation (17) yielded optimal solutions for planning horizons of 10, 100, and 1000 periods, respectively.

As a second example, when $\beta = 0.95$, $\rho = 0.7$, and $b/h = 25$, which corresponds to test problems 43, 44, and 45, we respectively obtained the following results:

- $S^* = 16$ and $\hat{S}^* = 18$, resulting in a convergence discrepancy of 4.10%.
- $S^* = 15$ and $\hat{S}^* = 14$, resulting in a convergence discrepancy of 1.52%.
- $S^* = 10$ and $\hat{S}^* = 11$, resulting in a convergence discrepancy of 0.40%.

Consequently, in test problems 43 to 45, Equation (17) yielded near-optimal solutions for planning horizons of 10, 100, and 1000 periods, respectively.

Theorem 2 suggests that the convergence discrepancy of Equation (17) tends to shrink and eventually diminishes as $T \rightarrow \infty$. This was indeed the case in the above two examples. However, note in Table 2 that there are a few special situations. For example, from case 7 to case 8 the convergence discrepancy increased (which is unexpected), whereas at the same time, from case 8 to case 9 the convergence discrepancy decreased (which is expected). Multiple forces could be at play here: First, the asymptotic convergence to a normal distribution as $T$ increases (which should in theory reduce the discrepancy) might not be smooth. Second, the increased impact of the integrality of the base-stock level due to the decrease of $S^*$, as $T$ increases, may increase the discrepancy.

Third, higher-order effects on the total cost distribution as $T$ increases, which is quite likely since the total cost distribution does contain some higher-order fluctuations as shown in Figure 1. The good news is that a review of Table 2 suggests that Equation (17) should produce optimal or near-optimal solutions in a wide variety of situations.

5.2  | Behavior of the total cost CVaR approximation as $T$ increases

Having concluded in Section 5.1 that Theorem 2 and Equation (17) can be trusted to produce solutions which are optimal or near-optimal in a wide variety of cases, we used Theorem 2 and Equation (17) in a second numerical experiment to study the behavior of $\hat{\zeta}_T(S, T)$ as $T$ increases.

In Figure 4 we present the second numerical experiment results. Figure 4 shows that for any $S$, $\hat{\zeta}_T(S, T) / T$ (total cost CVaR per period) is decreasing in $T$, even though $\hat{\zeta}_T(S, T)$ is obviously increasing in $T$.

As an additional observation, we know from Theorem 1 that $\zeta_T(S)$ is convex in $S$, but is the approximation $\hat{\zeta}_T(S)$ convex in $S$, too? Graphs A and C in Figure 4 show that $\hat{\zeta}_T(S)$ may be quasi-convex in $S$ and unimodal, while graphs B and D show that $\hat{\zeta}_T(S)$ may initially behave in a concave way, but after an inflection point will exhibit convex behavior.

6  | SENSITIVITY ANALYSIS

In this section, we explore the impact on $\hat{S}^*$ of changing the values of the various parameters. To increase the generality of our analysis, when appropriate, we will compare risk-averse firms interested in minimizing total cost CVaR with risk-neutral firms interested in minimizing expected total cost. We conducted a third numerical experiment in which we solve 5,400 instances of the problem under study using Theorem 2 and Equation (17). These instances resulted from all combinations of the following parameter values:

- $\mu = 1$
- $h = 1$
- $\beta = 0, 0.90, 0.95,$ and $0.99$ (4 values)
- $\rho = 0.30$ to $0.95$ in increments of $0.025$ (27 values)
- $b/h = 1$ to $50$ in increments of $1$ (50 values)
- $T = 10,000.$

We set $T = 10,000$ to make sure that the convergence discrepancy is negligible, and thus we maximize the probability that the total cost CVaR approximation yields optimal base-stock levels.

All $\hat{S}^*$ values and their corresponding $\hat{\zeta}_T(\hat{S}^*)$ values generated in this numerical experiment were stored in a database and are available from the authors upon request. The $\hat{S}^*$ values when $\beta = 0$ were obtained from Equation (11) in...
Figure 4: Examples of the approximate per-period total cost CVaR as a function of the base-stock level for several planning horizon lengths [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 5: Impact of $\beta$ and $\rho$ on $\hat{S}^*$ ($b/h = 50$)

Figure 6: Impact of $\beta$ and $b/h$ on $\hat{S}^*$ ($\rho = 0.9$)

Proposition 2: the $\hat{S}^*$ values when $\beta > 0$ were obtained from Theorem 2 and Equation (17). We created Figures 5 and 6 from representative subsets of the database. Furthermore, we prepared Tables 3 and 4 to respectively illustrate changes in $\hat{S}^*$ as $T$ increases and changes in $\hat{S}^*$ as $\beta$ increases.

Several observations can be derived from Figures 5 and 6 and Tables 3 and 4. We will loosely use the word “increase” to mean nondecrease because sometimes the optimal base-stock level will stay the same due to its integrality.

- **Observation 1.** As a function of capacity utilization and all else unchanged (ceteris paribus), the $\hat{S}^*$ of risk-averse and risk-neutral firms is increasing and increases at an increasing rate. However, the $\hat{S}^*$ of risk-averse firms increases as fast as or faster than that of risk-neutral firms. This observation can be verified in Figure 5. Observation 1 supports the intuitive notion that $\hat{S}^*$ increases in capacity utilization. Whereas Observation 1 confirms a similar result in Arreola-Risa (1996) for risk-neutral firms, to the best of our knowledge, Observation 1 represents an extension to the current operations management literature for risk-averse firms.
In addition, since risk-neutral firms are only concerned with the first moment of a total cost probability distribution and risk-averse firms are concerned with the whole total cost probability distribution, we find it intriguing that the $\hat{S}^*$ of both risk-neutral and risk-averse firms increases in a similar fashion as capacity utilization increases.

- **Observation 2.** As a function of the $bh/h$ ratio and all else unchanged, the $\hat{S}^*$ of risk-averse and risk-neutral firms is increasing and increases at a decreasing rate. However, the $\hat{S}^*$ of risk-averse firms increases as fast as or faster than that of risk-neutral firms.

This observation can be verified in Figure 6. Observation 2 supports the intuitive notion that $\hat{S}^*$ increases in the $bh/h$ ratio. Observation 2 represents an extension to the operations management research literature. Moreover, as mentioned in Observation 1, since risk-neutral firms are only concerned with the first moment of a total cost probability distribution and risk-averse firms are concerned with the whole total cost probability distribution, we find it intriguing that the $\hat{S}^*$ of both risk-neutral and risk-averse firms increases in a similar fashion as the $bh/h$ ratio increases.

- **Observation 3.** As a function of the planning horizon and all else unchanged, the $\hat{S}^*$ of risk-averse firms is nonincreasing and decreases at a decreasing rate. This observation can be verified in Table 3. We believe that Observation 3 is a new contribution to the operations management research literature. Observation 3 supports the intuitive notion that $\hat{S}^*$ will reach a maximum value as the planning horizon length approaches infinity.

- **Observation 4.** As a function of risk-aversion and all else unchanged, $\hat{S}^*$ is nondecreasing and increases at a decreasing rate. This observation can be verified in Table 4. We believe that Observation 4 is a new contribution to the operations management research literature. Observation 4 supports the intuitive notion that $\hat{S}^*$ will reach a maximum value as risk aversion grows without limit.

The combination of Observation 1 and Figure 5 yields a crucial discovery, from now on called **Finding 1**.

**Finding 1.** For specific values of $bh/h$, $\beta$, and $T$, say $b_0/h_0$, $\beta_0$, and $T_0$, there exists a threshold $r(\rho_0, \beta_0)$ such that $\hat{S}^* = S^*_g$ for any $\rho \leq \rho_0$ and $\beta \leq \beta_0$, where $\rho_0 > 0$ and $\beta_0 \in (0, 1)$.

The implication of Finding 1 is the following: Risk-averse firms with a $\beta$ in the range $0 < \beta \leq \beta_0$, a ratio $bh/h_0$, and a planning horizon $T_0$, which operate at a capacity utilization that does not exceed $\rho_0$, can simultaneously minimize total cost CVaR and expected cost by selecting a base-stock level equal to $S^*_g$. As an example of Finding 1, using Theorem 2 and Equation (17), one can determine for all risk-averse firms with parameters $bh/h = 50$, $T = 1000$, $0 < \beta \leq 0.95$ and $\rho \leq 54\%$, $\hat{S}^* = S^*_g = 6$ which would minimize both total cost CVaR and expected cost.

On a related matter, the combination of Observations 2-3 and Figure 6 leads us to a second crucial discovery, from now on called **Finding 2**.

**Finding 2.** Holding all else constant, since the $\hat{S}^*$ of risk-averse firms increases at a decreasing rate in $bh/h$ and

### TABLE 3 \( \hat{S}^* \) values for increasing values of $T$

| $\beta$ | $\rho$ | $bh/h$ | $T = 10$ | $T = 100$ | $T = 1000$ |
| --- | --- | --- | --- | --- | --- |
| 0.90 | 0.5 | 1 | 1 | 1 | 1 |
| | 5 | 3 | 2 | 2 | 2 |
| | 25 | 7 | 5 | 5 | 5 |
| 0.7 | 1 | 3 | 2 | 2 | 2 |
| | 5 | 9 | 7 | 5 | 5 |
| | 25 | 18 | 13 | 10 | 10 |
| 0.9 | 1 | 16 | 11 | 8 | 8 |
| | 5 | 38 | 30 | 21 | 21 |
| | 25 | 69 | 56 | 42 | 42 |
| 0.95 | 0.5 | 1 | 1 | 1 | 1 |
| | 5 | 4 | 3 | 2 | 2 |
| | 25 | 7 | 5 | 5 | 5 |
| 0.7 | 1 | 4 | 2 | 2 | 2 |
| | 5 | 10 | 7 | 5 | 5 |
| | 25 | 18 | 14 | 11 | 11 |
| 0.9 | 1 | 16 | 12 | 8 | 8 |
| | 5 | 39 | 31 | 22 | 22 |
| | 25 | 71 | 58 | 44 | 44 |

### TABLE 4 \( \hat{S}^* \) values for increasing values of $\beta$

| $\rho$ | $bh/h$ | $\beta = 0$ | $\beta = 0.5$ | $\beta = 0.95$ |
| --- | --- | --- | --- | --- |
| 0.5 | 10 | 3 | 4 | 5 |
| | 30 | 4 | 6 | 7 |
| | 50 | 5 | 7 | 8 |
| 0.6 | 10 | 4 | 5 | 7 |
| | 30 | 6 | 8 | 11 |
| | 50 | 7 | 10 | 13 |
| 0.7 | 10 | 6 | 9 | 12 |
| | 30 | 9 | 14 | 17 |
| | 50 | 11 | 16 | 20 |
| 0.8 | 10 | 10 | 17 | 21 |
| | 30 | 15 | 25 | 31 |
| | 50 | 17 | 29 | 36 |
| 0.9 | 10 | 22 | 43 | 51 |
| | 30 | 32 | 64 | 74 |
| | 50 | 37 | 74 | 85 |
decreases at a decreasing rate in \( T \), there are ranges of \( b/h \) values and \( T \) values for which \( \hat{S} \) varies little, if at all.

As an example of Finding 2, consider a collection of risk-averse firms having all combinations of the following parameter values: \( \beta = 0.95, \rho = 0.7, b/h \in [80,100], \) and \( T \in [3000, 15000] \). Then using Theorem 2 and Equation (17), we find that \( \hat{S} = 6 \) for this collection of risk-averse firms. As a second example of Finding 2, consider a collection of risk-averse firms having all combinations of the following parameter values: \( \beta = 0.95, \rho = 0.5, b/h \in [80,100], \) and \( T \in [3000, 15000] \). Then using Theorem 2 and Equation (17), we obtain \( \hat{S} = 6 \) for this collection of risk-averse firms.

7 | DISCUSSION AND APPLICATION OF RESULTS

In this section, we detail how some of our research results, observations, findings, and insights were applied in the supply-chain design project of Company Delta. Unconventional oil is petroleum that cannot be extracted by conventional drilling methods. Energy companies worldwide have been investing in unconventional oil technologies due to the increasing scarcity of conventional oil reserves. Over several decades, the R&D division of Company Delta has developed a successful technology for extracting unconventional oil. Such technology entails the use of a complex, heavy, voluminous, and extremely expensive apparatus which is deployed in the ground for petroleum extraction. For brevity, in this paper the apparatus will be referred to as Product Z. The group of workers in charge of deploying one unit of Product Z will be called a deployment crew.

Company Delta wants to take Product Z from R&D to large-scale production. For this purpose, Company Delta launched a project to design a supply chain that would make large-scale production of Product Z financially attractive. The starting point of the project was to focus on the last two stages of the supply chain, namely, final assembly and deployment. These stages could be decoupled from the rest of the supply chain because all raw materials and components would be provided to final assembly in a just-in-time fashion. Two of the three major supply chain design strategic considerations were the long-run average deployment rate and the long-run average production rate in the final assembly facility. These two average rates were denoted by \( \lambda \) and \( \mu \), respectively. The third major supply chain design strategic consideration was planning horizon length, which was denoted by \( T \). A trio of \(( \lambda, \mu, T )\) values was called a supply-chain design.

To buffer the random interaction of deployment and production, Company Delta would maintain an inventory of Product Z. Because ordering costs were negligible, such buffer inventory would be managed by a base-stock policy: Starting from a base-stock \( S \), every unit demanded by a deployment crew would trigger a production order at the final assembly facility. Note that a unit demanded by a deployment crew when the inventory of Product Z was temporarily depleted implies that the demand for the unit would be naturally back-ordered; that is, the deployment crew would have to wait until a unit of Product Z became available to them. In addition, note that the cost of having a deployment crew idle while waiting for a unit of Product Z to deploy is equivalent to a back-ordering cost per unit per period. At the same time, units in inventory would incur a holding cost measured in dollars per unit per period. Let total cost be defined as the sum of the inventory-holding and back-ordering costs that the supply chain will incur over the planning horizon length \( T \). Because the order of magnitude of the total cost in the supply-chain design project was estimated to be in the billions of dollars, and the occurrence of total costs in the right tail of the distribution would be financially catastrophic, due to its risk aversion, Company Delta decided to focus on creating a supply-chain design which would minimize total cost CVaR.

7.1 | Product Z final assembly

The manufacturing facility in which Product Z would be assembled was to be built near the deployment site. By design, the manufacturing facility would assemble Product Z one unit at a time and operate 24/7. Product Z would be assembled from four major components which in this paper will be called I, II, III, and IV. Components I and II weighed 6 tons each, component III weighed 3 tons, and component IV weighed 2 tons. The final position of component III is inside component I, and the final position of component IV is inside component II. Once component III was inside component I and component IV was inside component II, component III was to be joined to component IV, and then component I was to be joined to component II. The length of these components was in the order of magnitude of thousands of feet. Due to the weight and length of the components in combination with the insertion and joining processes, it was clear to the engineering firm in charge of designing the manufacturing facility that unit production times would be random. Let \( T_p \) denote the random time in hours to assemble one unit and \( P \) denote the final assembly rate in units per day. Then, as shown in Appendix 10, by a straightforward application of the elementary renewal theorem, we know that the long-run average daily assembly rate \( \mu \) is given by

\[
\mu \equiv \frac{E(P)}{24/E(T_p)}.
\] (18)

Although the manufacturing facility was being designed to yield a target value of \( \mu \), the probability density function of \( T_p \) was impossible to specify in advance. Thus, to model \( T_p \) we used the principle of maximum entropy which says the following (see eg. Kapur & Kesavan, 1992): If nothing is known about a probability distribution except its mean, then the probability distribution with the largest entropy should be chosen because by maximizing entropy, the amount of prior information built into the probability distribution is minimized, and at the same time, many physical systems tend to move towards
maximal entropy configurations over time. Since we know that \( T_p \) is a continuous positive random variable with mean \( 24/\mu \) and the exponential distribution is the maximum entropy distribution among all continuous positive distributions that have a specified mean (Jaynes, 1957a, 1957b), we decided to model \( T_p \) as an exponential random variable with mean \( 24/\mu \).

### 7.2 Product Z deployment

The random time in hours to deploy one unit by one deployment crew is defined as \( T_d \), and \( \mathbb{E}(T_d) \) as its expected value. Based on preliminary tests, Company Delta was able to estimate the value of \( \mathbb{E}(T_d) \) and obtain the probability density function of \( T_d \), from now on denoted by \( f_{T_d}(\cdot) \). Interestingly enough, \( f_{T_d}(\cdot) \) did not resemble any of the well-known probability density functions. Let \( C \) denote the number of deployment crews working concurrently and \( \lambda \) denote their long-run average daily deployment rate. Given that the deployment crews will work 24/7, then, as shown in Appendix 11, by a straightforward application of the elementary renewal theorem, we know that

\[
\lambda = 24C/\mathbb{E}(T_d).
\]

For instance, if one deployment crew can deploy one unit in 48 hours on average (\( \mathbb{E}(T_d) = 48 \)), and four deployment crews are working concurrently (\( C = 4 \)), then their long-run average deployment rate would be \( \lambda = 2 \) units per day. Note that \( \lambda \) may also be interpreted as the average demand per day of units to be deployed by the \( C \) deployment crews. In order to achieve higher or lower values of \( \lambda \), \( C \) would be increased or decreased accordingly.

Let \( T_{d1}^1, T_{d2}^2, T_{d3}^3, \ldots \) be the sequence of IID deployment epochs by one deployment crew. Then the stochastic process which counts the number of deployments over time by one deployment crew is a renewal process. In consequence, the random deployment rate \( D \) of the combined \( C \) deployment crews is equal to the superposition of \( C \) renewal processes. Additionally, based on the estimated value of \( \mathbb{E}(T_d) \) and the target values of \( \lambda \), we knew that \( C \) was going to be a very large number (which cannot be disclosed in this paper due to confidentiality).

Therefore, armed with the well-known result regarding the superposition of a large number of renewal processes converging to a Poisson process (see, eg Cinar & Agnew, 1968), we decided to simulate \( D \) using the \( f_{T_d}(\cdot) \) obtained in the preliminary tests in combination with the range of values of \( C \) under consideration. The simulation results showed that a Poisson process with parameter \( \lambda \) was indeed a very accurate model of \( D \). Thus we decided to model \( D \) as a Poisson process with parameter \( \lambda \).

### 7.3 Product Z inventory policy

As mentioned earlier, Company Delta decided to keep an inventory of Product Z as a buffer between the randomness in the deployment rate \( D \) and the randomness in the final assembly rate \( P \). Let \( h \) denote the inventory-holding cost in dollars per unit per day. Furthermore, given that set-up costs were nonexistent and ordering costs were negligible, the following base-stock policy was selected to manage the inventory.

(a) Starting with \( S \) units on hand of Product Z, each unit of Product Z demanded by a deployment crew would trigger a production order for one unit at the final assembly facility. The initial \( S \) units of Product Z would be produced during a ramp-up period and before deployment starts.

(b) If there were units of Product Z on hand when a deployment crew was ready to deploy a unit, then the deployment crew would get the unit and proceed to deploy it.

(c) If there were no units of Product Z on hand when a deployment crew was ready to deploy a unit, then the deployment crew would wait for a unit to become available, and a delayed deployment cost (or demand back-ordering cost) of \( b \) dollars per day would be incurred.

### 7.4 Product Z economic parameters

The production cost of one unit of Product Z was denoted by \( c \). The value of \( c \) depended on the final design of Product Z and on the prices of several metals such as carbon steel, stainless steel, and copper. Due to the volatility of these prices and the various designs of Product Z being tested, a wide range of \( c \) values was being considered. At the same time, \( r \) be the daily inventory holding cost expressed as a percentage. Because the value of \( r \) among other factors reflected financial risk, a wide range of \( r \) values was also being considered. Consequently, since \( h = c \cdot r \), the range of resulting \( h \) values was even wider than the individual ranges of the \( c \) and \( r \) values.

At the same time, the values of the delayed deployment cost \( b \) depended on the opportunity cost of having a deployment crew idle, as well as on oil prices, which were inherently volatile, and on several other economic variables. For that reason, a wide range of \( b \) values was to be included in the determination of the best supply-chain design.

### 7.5 Creating an optimal supply-chain design

The volatility of the economic parameters \( c, r, b \) led Company Delta to adopt a scenario-driven approach to assess many different supply-chain designs \( (\lambda, \mu, T) \) in order to arrive at the optimal one. To the extent allowed by our confidentiality agreement with Company Delta, we will explain in this section how some of the results in this paper guided Company Delta in the search for the optimal supply-chain design. The numerical information we use in the explanations will be realistic but not real.
First, Theorem 2 and Equation (17) were used to compare numerous supply-chain designs under copious scenarios. This was a tremendous time-and-money saver, especially when compared to brute-force simulation. Of course, once a particular supply-chain design was identified as a candidate for the best design, its optimal base-stock level and corresponding total CVaR obtained from Theorem 2 and Equation (17) would be verified via simulation.

Second, Observations 1 and 2 were used to make predictions regarding the optimal base-level behavior as supply-chain designs were being generated. These observations were also helpful to better understand the trade-offs among the various supply-chain designs under consideration. Moreover, Observation 3 was instrumental in understanding the impact of planning horizon length on the optimal base-stock level for the various supply-chain designs when compared across many different scenarios.

Third, without question, the combination of Findings 1 and 2 revolutionized the way in which Company Delta approached the creation of the optimal supply chain design. To explain why, let us assume that Company Delta was considering a $\beta = 0.95$, $T \in [4000, 8000]$ days, $c \in \{2200, 2260, 22600\}$, and $b \in \{2965, 3000\}$ per unit per day. Furthermore, let us assume that to account for financial risk, Company Delta was considering that the annual inventory holding cost was between 5% and 6%. Assuming 365 days per year and that daily compounding is negligible leads to $r \in [0.0137\%, 0.0164\%]$. Then the range of $h = r \cdot c$ can be calculated to be $\{30, 37\}$ per unit per day. In turn, the range of $b/h$ can be calculated to be $[80, 100]$.

Based on Theorem 2, Equation (17) and Findings 1 and 2, Company Delta could ascertain that under the large number of scenarios resulting from taking all combinations of $T \in [4000, 8000]$ days and $b/h \in [80,100]$, for any value of $\lambda$ being considered,

(a) if the value of $\mu$ was selected such that $\rho = 80\%$, then $\hat{S}^* \in [21, 23]$;
(b) if the value of $\mu$ was selected such that $\rho = 75\%$, then $\hat{S}^* \in [16, 17]$;
(c) if the value of $\mu$ was selected such that $\rho = 70\%$, then $\hat{S}^* \in [12, 13]$;
(d) if the value of $\mu$ was selected such that $\rho \geq \rho_0 = 67\%$, then $\hat{S}^* = S^*_2 = 11$ and this base-stock level would simultaneously minimize expected total cost and total cost CVaR.

Knowing, for example, that $\hat{S}^* \in [21, 23]$ when $\rho = 80\%$, $T \in [4000, 8000]$ days, and $b/h \in [80,100]$, greatly simplified the creation of an optimal supply-chain design. More importantly, the realization that it was possible to simultaneously minimize total cost CVaR and expected total cost (eg, in part (d) above) became a tipping point in the creation of the optimal supply-chain design because that realization induced Company Delta to consider values of $\rho$ that otherwise would have been ignored.

We will close this section by mentioning that how Company Delta specifically created the optimal supply-chain design is not only confidential, but beyond the scope of this paper as well. However, it should now be obvious to the reader that our research results played a key role in the creation of the optimal supply-chain design.

8 | CONCLUSION

We addressed the problem of finding the base-stock level that minimizes the total cost CVaR of a single-product, make-to-stock, risk-averse firm which uses a base-stock policy to control inventory and production. Product demand and unit manufacturing time are stochastic and were respectively modeled as a Poisson process and an exponential random variable. Stock-outs are back-ordered, and total cost is equal to the sum of inventory-holding and back-ordering costs over a continuous-time, finite planning horizon.

Our research results fall at the intersection of two major and independent operations management research streams. The first research stream minimizes downside risk in firms which either purchase the product from suppliers, produce the product in a manufacturing system with infinite capacity, and/or make a single inventory-production decision at the beginning of a one-period planning horizon. The second research stream minimizes expected cost in firms which have either a capacitated make-to-stock system or multiple inventory-production decisions made over a finite planning horizon, or both. By extending the two research streams, our results make it possible to capture the inherent trade-offs among risk aversion, inventory, manufacturing capacity, and planning horizon length when aiming to minimize downside risk.

This article has made four contributions. First, we demonstrated that total cost CVaR is convex in the base-stock level, increasing in risk aversion, and jointly convex in per-period unit inventory-holding and per-period unit back-ordering cost. Second, we derived closed-form approximations of expected cost, variance of total cost, and total cost CVaR, which are all asymptotically exact in planning horizon length. The approximations yielded optimal or near-optimal base-stock levels in a wide variety of test problems. Third, we articulated observations and findings regarding the impact on the optimal base-stock levels produced by the approximations of changes in risk aversion, manufacturing capacity, per-period unit inventory-holding cost, per-period unit back-ordering cost, and planning horizon length. Fourth, we provided a detailed description of how our research results were applied in a real-world supply-chain design project.

Several avenues may be pursued to extend the preliminary results of this paper. Although variance-at-risk (VaR) is not a coherent risk measure and does not control scenarios...
exceeding the VaR value, given its popularity in downside risk literature, it would be of theoretical and practical interest to consider the problem of finding the base-stock level that minimizes total cost VaR in a setting similar to ours. A second research avenue would be to study the problem of finding the base-stock level that minimizes the expected total cost subject to a maximum value of total cost CVaR in a setting similar to ours. We believe that Propositions 1 and 3 as well as Theorem 2 should provide an excellent starting point for exploring these two research avenues.

A third research avenue may be to model demand as a renewal process and to model unit production time by any of the known probability distributions such as gamma, normal, or uniform. A perusal of the stochastic processes and queuing theory literature suggests that in this research avenue, it may not be possible to find tractable expressions such as those in Proposition 1, Proposition 3, and Theorem 2. Nevertheless, the ingenuity of our fellow researchers may prove us wrong.

Additional research avenues would be consideration of multiple products, multiple production stages, or both, as well as consideration of economic dimensions such as pricing or allowing for manufacturing capacity expansions at a cost.

ACKNOWLEDGMENTS

We thank the reviewers and the editors for their suggestions that significantly improve this paper. Bo Li acknowledges financial support from Programs Supported by Ningbo Natural Science Foundation (Grant nos. 2018A610120 and 2019A610047), Zhejiang Qianjiang Talent Program (Grant no. QJC1803003), and the NISCI-TDH Line-haul Road Transportation Lab.

ORCID

Bo Li https://orcid.org/0000-0003-4490-5402

REFERENCES

Amihud, Y., & Lev, B. (1981). Risk reduction as a managerial motive for conglomerate mergers. The Bell Journal of Economics, 12, 605–617.

Arreola-Risa, A. (1996). Integrated multi-item production-inventory systems. European Journal of Operational Research, 89(2), 326–340.

Borgonovo, E., & Peccati, L. (2009). Financial management in inventory problems: Risk averses vs risk neutral policies. International Journal of Production Economics, 118(1), 233–242.

Carneiro, M. C., Ribas, G. P., & Hamacher, S. (2010). Risk management in the oil supply chain: A CVaR approach. Industrial & Engineering Chemistry Research, 49(7), 3286–3294.

Chen, X., Shum, S., & Simchi-Levi, D. (2014). Stable and coordinating contracts for a supply chain with multiple risk-averse suppliers. Production and Operations Management, 23(3), 379–392.

Chen, Y. F., Xu, M., & Zhang, Z. G. (2009). Technical note – A risk-averse newsvendor model under the CVaR criterion. Operations Research, 57(4), 1040–1044.

Cinlar, E., & Agnew, R. (1968). On the superposition of point processes. Journal of the Royal Statistical Society. Series B (Methodological), 30(3), 576–581. https://doi.org/10.1111/j.2517-6161.1968.tb00758.x.

Dave, C., Eckel, C. C., Johnson, C. A., & Rojas, C. (2010). Eliciting risk preferences: When is simple better? Journal of Risk and Uncertainty, 41(3), 219–243.

DeCroix, G. A., & Arreola-Risa, A. (1998). Optimal production and inventory policy for multiple products under resource constraints. Management Science, 44(7), 950–961.

Gotoh, J.-Y., & Takano, Y. (2007). Newsvendor solutions via conditional value-at-risk minimization. European Journal of Operational Research, 179(1), 80–96.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. Physics Review, 106, 620–630.

Jaynes, E. T. (1957b). Information theory and statistical mechanics. II. Physics Review, 108, 171–190.

Kaplan, R. S. (1970). A dynamic inventory model with stochastic lead times. Management Science, 16(7), 491–507.

Kapur, J. N., & Kesavan, H. K. (1992). Information theory and statistical mechanics. II. Physics Review, 108, 171–190.

Kaplan, R. S. (1970). A dynamic inventory model with stochastic lead times. Management Science, 16(7), 491–507.

Karmarkar, U. S. (1987). Lot sizes, lead times and in-process inventories. Management Science, 33(3), 409–418.

Lee, Y.-J., & Zipkin, P. (1992). Tandem queues with planned inventories. Operations Research, 40(5), 936–947.

Lee, Y.-J., & Zipkin, P. (1995). Processing networks with inventories: Sequential refinement systems. Operations Research, 43(6), 1025–1036.

Li, B., & Arreola-Risa, A. (2017). Financial risk, inventory decision and process improvement for a firm with random capacity. European Journal of Operational Research, 260(1), 183–194.

Rao, S., & Goldsby, T. J. (2009). Supply chain risks: A review and typology. International Journal of Logistics Management, 20(1), 97–123.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. Journal of Banking & Finance, 26(7), 1443–1471.

Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. Management Science, 46(3), 404–420.

Serrano, A., Oliva, R., & Kraiselburd, S. (2017). On the cost of capital in inventory models with deterministic demand. International Journal of Production Economics, 183, 14–20.

Sladký, K., & van Dijk, N. M. (2005). Total reward variance in discrete and continuous time Markov chains. In Operations research proceedings 2004 (pp. 319–326). Berlin, Heidelberg: Springer.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. Science, 315(5811), 515–518.

Tomlin, B., & Wang, Y. (2005). On the value of mix flexibility and dual sourcing in unreliable newsvendor networks. Manufacturing & Service Operations Management, 7(1), 37–57.

U.S. Board of Governors of the Federal Reserve System. (2019). Capacity utilization: Total industry [TCU]. https://fred.stlouisfed.org/series/TCU

van Dijk, N. M., & Sladký, K. (2006). On the total reward variance for continuous-time Markov reward chains. Journal of Applied Probability, 43(4), 1044–1052.

Veinott, A. F. J. (1965). Optimal policy for a multi-product, dynamic, nonstationary inventory problem. Management Science, 12(3), 206–222.
Wagner, M. R., & Radovilsky, Z. (2012). Optimizing boat resources at the US Coast Guard: Deterministic and stochastic models. *Operations Research, 60*(5), 1035–1049.

Wu, M., Zhu, S. X., & Teunter, R. H. (2013). The risk-averse newsvendor problem with random capacity. *European Journal of Operational Research, 231*, 328–336.

Zhang, D., Xu, H., & Wu, Y. (2009). Single and multi-period optimal inventory control models with risk-averse constraints. *European Journal of Operational Research, 199*(2), 420–434.

Zipkin, P. H. (1986). Models for design and control of stochastic, multi-item batch production systems. *Operations Research, 34*(1), 91–104.

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Li B, Arreola-Risa A. On minimizing downside risk in make-to-stock, risk-averse firms. *Naval Research Logistics* 2021;68:199–213. https://doi.org/10.1002/nav.21941