Base-Stock Policies with Constant Lead Time: Closed-Form Solutions and Applications

Zhaolin Li

Discipline of Business Analytics, The University of Sydney Business School, Australia erick.li@sydney.edu.au

Guitian Liang

School of Management, Jinan University, Guangzhou, China lianggt@hotmail.com

Qi Fu

The University of Macau, Taipa, Macau, China, gracefu@um.edu.mo

Chung-Piaw Teo

NUS Business School, National University of Singapore, Singapore 119245, bizteocp@nus.edu.sg

We study stationary base-stock policies for multiperiod dynamic inventory systems with a constant lead time and independently and identically distributed (iid) demands. When ambiguities in the underlying demand distribution arise, we derive the robust optimal base-stock level in *closed forms* using only the mean and variance of the iid demands. This simple solution performs exceptionally well in numerical experiments, and has important applications for several classes of problems in Operations Management.

More important, we propose a new distribution-free method to derive robust solutions for multiperiod dynamic inventory systems. We formulate a zero-sum game in which the firm chooses a base-stock level to minimize its cost while Nature (which is the firm's opponent) chooses an iid two-point distribution to maximize the firm's time-average cost in the steady state. By characterizing the steady-state equilibrium, we demonstrate how lead time can affect the firm's equilibrium strategy (i.e., the firm's robust base-stock level), Nature's equilibrium strategy (i.e., the firm's most unfavorable distribution), and the value of the zero-sum game (i.e., the firm's optimized worst-case time-average cost). With either backorders or lost sales, our numerical study shows that superior performance can be obtained using our robust base-stock policies, which mitigate the consequence of distribution mis-specification.

Key words: Inventory management, robust optimization, closed form, zero-sum game

History: Updated on 24 February, 2023

1. Introduction

A positive lead time makes stockouts common in supply chains (Ergin et al. 2022) and costly for retailers (Corsten and Gruen 2004). In the inventory management literature, a system with lost sales and a positive lead time is widely regarded as notoriously difficult but fundamentally important. A technical hurdle arises when tracking the pipeline inventory where the state space rapidly expands as the lead time grows. The curse of dimensionality inevitably hinders the analysis. Numerous articles have proposed heuristics/algorithms to compute policies with satisfactory numerical performance; however, these methods are computationally intensive and cannot be easily implemented, e.g., in an Excel spreadsheet, which is a popular tool used by many practitioners. Furthermore, these methods rely on a known prior distribution fitted from the training data. The prior distribution could lead to model mis-specification and may not perform well if the demand distribution experiences a distributional shift.

Motivated by these theoretical and practical challenges, we apply the robust max-min decision rule in a multiperiod setting to develop closed-form solutions for base-stock policies in classical inventory problems. Our solutions are not only suitable for implementing in Excel spreadsheets but also deliver high quality numerical performance. We first consider the case of backorders and then extend the analysis to the case of lost sales. We assume that the lead time is a positive constant and cost parameters are linear and stationary. We also assume that the demand in each period is identically and independently distributed (iid) but with an *unknown* distribution. The partially available information includes the mean and variance of the iid demand. An innovative feature of our method is that we formulate the robust optimization model as a zero-sum game, in which the firm chooses a base-stock level to minimize her cost while Nature chooses an iid distribution to maximize the firm's time-average cost in the steady state. This method overcomes the technical hurdle encountered in the extant literature. Specifically, in the backorder model, we find that Nature's equilibrium strategy is a two-point distribution. For the lost-sales model, due to the difficulty in precisely deriving the objective function, we consider a relaxed model where Nature chooses only two-point distributions. We then characterize the (relaxed) zero-sum game equilibrium. Our numerical study shows that the firm's optimal base-stock level in the relaxed solution performs remarkably well.

1.1 Literature Review

In the extant literature, various heuristics have been developed for lost-sales models (see Goldberg et al. 2021, for an updated literature review). For example, Janakiraman and Roundy (2004) showed that with a random lead time, the firm's expected discounted total cost is convex with respect to the base-stock level. Zipkin (2008a) formulated a large-scale Markov chain to characterize the optimal policy, which is state-dependent. To ease the computational effort, Levi et al. (2008) introduced a dual-balancing policy to reduce the holding cost and lost-sales cost. Zipkin (2008b) showed that the optimal cost function exhibits the $L^{\#}$-convex (also known as L-natural convex) property, and Chen et al. (2014) extended the concept of $L^{\#}$-convexity to develop an algorithm with pseudopolynomial computational time. Huh et al. (2009) used backorder solutions to construct a weighted-average base-stock level for the lost-sales counterpart. By aggregating the dynamics of the inventory system into a one-dimensional state space, Arts et al. (2015) developed a new heuristic to choose the base-stock level. van Jaarsveld and Arts (2021) considered a projected inventory level policy, under which the expected inventory level is raised to a fixed level. Wei et al. (2021) solved a deterministic model with backorders and then showed that the resultant policy is asymptotically optimal with backorders or lost sales. Xin (2021) demonstrated that capped base-stock policies (which include a base-stock level and a cap on the order size) perform satisfactorily.

In addition to the base-stock policies, the constant order policy (COP), under which the firm orders the same amount of inventory in each period irrespective with the pipeline and on-hand inventory, has received increasing attention in the literature. Goldberg et al. (2016) showed that COP can be asymptotically optimal under a long lead time or high understock cost. Xin and Goldberg (2016) provided explicit bounds for the profit gap between COP and the optimal policy and then proved that this gap converges to zero with an exponential speed as the lead time increases. The subsequent studies focus on determining the COP quantity in various environments such as dual sourcing (Xin and Goldberg 2018) and random supply (Bu et al. 2020). In contrast, we consider stationary base-stock policies with unknown demand distributions.

The literature on robust inventory management is inspired by Scarf (1958), and single-period models are prevalent. A typical assumption is that the firm knows only the mean and variance of the demand. The firm aims to find an order quantity to maximize her expected profit against the worst possible distribution. While Scarf's research applied semi-infinite programming (SIP) tools, Gallego and Moon (1993) provided an alternative proof using the Cauchy inequality. These traditional methods focused on the firm's perspective and encountered a few roadblocks when the ex post profit function is complex or the t-th moment (where $t > 1$) is known. In contrast, Li et al. (2022) proposed a more efficient method to analyze the single-period model from Nature's perspective, thus making closed-form solutions attainable. Multiperiod robust inventory models are rare. Mamani et al. (2017) proposed a method that combines the central limit theorem with robust optimization tools. Xin and Goldberg (2022) developed a multiperiod model with zero lead time, backorders, and Martingale demand. The main difference is that Martingale demand provides partial information about the mean of the demand (which equals the realization of the most recent demand).

1.2 Our Contributions

We highlight our main contributions and results as follows.

- We propose a novel method to solve multiperiod robust inventory models as a zero-sum game. We consider the steady-state equilibrium, in which both the firm and Nature employ stationary strategies. We obtain the closed-form solution in the firm's relaxed model (where Nature chooses only two-point distributions). We show that the firm's relaxed solution is optimal for the general backorder model and a special case of the lost-sales model.

- We show that backorders and lost sales create different impacts on the relaxed equilibrium. With lost sales, the equilibrium strategy of Nature is unaffected by the lead time. The intuition is that with lost sales, the impact of shortage only lasts for one period (rather than accumulates), forcing Nature to resume the single-period strategy. Hence, the value of the zero-sum game is also unaffected by the lead time (but the base-stock level is affected by the lead time). In contrast, with backorders, the demand history is preserved, and accumulated shortages can grow, enabling Nature to employ a lead-time-dependent strategy to punish the firm. Both the firm's equilibrium strategy and the value of the zero-sum game depend on the lead time under backorders. Therefore, backorders increase the firm's cost.

- Our numerical study shows that the robust base-stock policy generated from the relaxed model performs satisfactorily. Due to the notorious difficulty in using the exact demand distribution to compute the optimal policy for the lost-sales models, our closed-form robust solution is an attractive alternative with significant practical and theoretical value. This simple but elegant equation is suitable for implementation in Excel spreadsheets and comprehensively captures the effect of lead time.

2. The Backorder Model with Constant Lead Time

2.1 Known Demand Distribution

In the model with backorders, the planning horizon is divided into multiple discrete periods, of which the lengths are identical and scaled to one unit of time. Any replenishment order arrives after a constant lead time l (where $l \geq 1$ is an integer). The holding cost for on-hand inventories is \$$h$ and the backorder cost is \$$b$ per unit per period. In each period, the same sequence of events proceeds as follows: (1) the previously scheduled shipment arrives; (2) the firm places an order; (3) the external demand for the current period is realized; and (4) holding and backorder costs are incurred. Hereafter, we use a tilde to indicate a random variable and no tilde to indicate a realized value. The external demand that the firm receives in each period is represented by a nonnegative random variable $\tilde{\theta}$, which follows an iid cumulative distribution function F.

The literature has demonstrated that the optimal inventory policy when the firm precisely knows the distribution F is a *stationary base-stock policy* (e.g., see Section 2.1 of Shang and Song 2003, for details). We let s be the base-stock inventory level and $\xi = \sum_{i=1}^{l+1} \theta_i$ be the realized demand during the lead time and the current period. In the steady state, the time-average cost equals

$$M(s|F) = E\left[b\left(\tilde{\xi} - s\right)^{+} + h\left(s - \tilde{\xi}\right)^{+}\right], \tag{2.1}$$

which depends on the iid demand distribution F. The unambiguous optimal base-stock level satisfies $F_{l+1}(\hat{s}) = \frac{b}{b+h}$, where $F_{l+1}(\cdot)$ is the $(l+1)$-fold convolution of $F(\cdot)$.

2.2 Unknown Demand Distribution

We assume that the exact form of F is unknown except that the values of the mean ($E(\tilde{\theta}) = \mu > 0$) and variance ($Var(\tilde{\theta}) = \sigma^2 > 0$) are known. Let Ω be the collection of all the possible distributions satisfying the mean and variance conditions. Formally, we define set Ω as follows:

$$\Omega = \left\{F(\cdot) \,|\, \int_0^\infty dF(\theta) = 1, \int_0^\infty \theta dF(\theta) = \mu, \int_0^\infty \theta^2 dF(\theta) = \mu^2 + \sigma^2\right\}. \tag{2.2}$$

The number of feasible distributions in set Ω is infinite and each feasible distribution satisfies the mean-variance constraints. We let $\rho = \frac{\sigma}{\mu}$ be the coefficient of variation of the demand. Under the assumption of iid demands, the mean and variance of $\tilde{\xi}$ satisfy $E(\tilde{\xi}) = (l+1)\,\mu$ and $Var(\tilde{\xi}) = (l+1)\,\sigma^2$, respectively. Because the stationary base-stock policy is optimal for the unambiguous benchmark, we assume that the firm uses a stationary base-stock policy when F is unknown. The firm applies the robust max-min decision rule to solve the following robust optimization model in the *steady state*:

$$Z = \min_{s\geq 0} \sup_{F\in\Omega} \{M\,(s|F)\} = \min_{s\geq 0} \sup_{F\in\Omega} \left\{ \int_0^{\infty} b\,(\xi - s)^+ + h\,(s-\xi)^+ \, dF_{l+1}(\xi) \right\}, \tag{2.3}$$

which has a neat zero-sum game interpretation. Note that for any distribution $F_0 \in \Omega$, we have $\sup_{F\in\Omega} \{M\,(s|F)\} \geq M\,(s|F_0)$; hence Z obtained above is an upper bound to the optimal cost $\min_{s\geq 0} M\,(s|F_0)$, when the iid demand distribution F_0 is known.

We introduce Nature as the second player of the game. The firm chooses an inventory level s to minimize the expected cost $M\,(s|F)$ while Nature chooses a distribution F to maximize the firm's expected cost $M\,(s|F)$. The set Ω in equation (2.2) specifies Nature's strategy space. Denote the equilibrium of this zero-sum game by (s^*, F^*), where s^* is the firm's equilibrium strategy (or her distributionally robust inventory level) and F^* is Nature's equilibrium strategy (or the firm's most unfavorable distribution). Consequently, $Z^* = M\,(s^*|F^*)$ represents the value of the zero-sum game (or the firm's optimized worst-case expected cost). We clarify that when the firm plays an arbitrary inventory level s, Nature responds by playing its best response $F_{bst}\,(s)$. Evidently, it holds that $F_{bst}\,(s^*) = F^*$ and for any other $s \neq s^*$, $F_{bst}\,(s) \neq F^*$, indicating that the equilibrium strategy of Nature differs from its best response to an arbitrary base-stock level s. We focus on deriving the equilibrium strategies (F^*, s^*) rather than the off-equilibrium actions. To avoid uninteresting results, we assume $b \geq \rho^2 h$ (which is consistent with Scarf's rule) throughout this paper. Otherwise, the firm may backorder forever without keeping any on-hand inventory.

2.3 A Relaxed Solution

Conceptually, the firm and Nature engage in a multiperiod zero-sum game. The popular methods to solve the game include the subgame perfect conditions and the one-shot deviation principle (Fudenberg and Tirole 1991). Because we are interested in the steady-state equilibrium where both the firm and Nature employ stationary strategies, these two popular methods become less effective. For example, with a positive lead time (and later with lost sales), the number of subgames grows rapidly. Under the iid assumption, if Nature deviates from its initial strategy in one shot, it must keep playing the same deviating strategy in every period and cannot resume to its initial strategy. Therefore, whenever Nature deviates from its initial strategy, we must examine the changes over the entire planning horizon rather than just one short period. To overcome these challenges, we first consider a relaxed model for the firm.

To understand the motivation for this relaxation, we note that in (2.3), $\xi = X_1 + \ldots + X_{l+1}$, where X_i are iid random variables with prescribed mean and variance, and the optimal solution depends on the critical fractile at $b/(b+h)$ for ξ. To simplify the problem, we use the fact that any real-valued random variable X_i has a Bernoulli decomposition in the form $X_i \sim Y_p(t_i) + \delta_p(t_i)\eta_i$ where t_i is uniformly distributed in $(0,1)$, η_i is independent of t_i and is a Bernoulli random variable taking value in 0 or 1, with probability $1-p$ and p respectively (cf. Aizenman et al. 2022). In particular, if we choose $p = (b/(b+h))^{1/(l+1)}$, then $\tilde{\xi} = \sum_{i=1}^{l+1} Y_p(t_i)$ with probability $b/(b+h)$, which is the critical ratio in the problem. The challenge is to characterize how this critical fractile changes as each t_i varies over $(0,1)$ independently in a uniform manner.

Note that each X_i can be represented as mixture of Bernoulli random variables. In the rest of the paper, we restrict the problem to the space of Bernoulli distributions to derive the equilibrium solution. Our analysis shows that the worst case distribution for X_i in this setting is actually a Bernoulli random variable, with $p = (b/(b+h))^{1/(l+1)}$, and values chosen to satisfy the mean and variance constraints.

Let Ω_0 be the collection of two-point distributions satisfying the following equation:

$$\begin{cases} \Pr\left(\tilde{\theta} = \mu - \sigma\sqrt{\frac{1-\beta}{\beta}} \overset{def}{=} L\right) = \beta, \\ \Pr\left(\tilde{\theta} = \mu + \sigma\sqrt{\frac{\beta}{1-\beta}} \overset{def}{=} H\right) = 1-\beta. \end{cases} \tag{2.4}$$

In this relaxed model, Nature chooses the demand distribution F from Ω_0, defined by equation (2.4). We omit the nonnegative constraint on $\tilde{\theta}$ but retrospectively verify that this constraint still holds in the equilibrium. Interestingly, due to the zero-sum feature of the game, Nature's *constrained* model becomes the firm's *relaxed* model. When Nature's strategy space shrinks from Ω to Ω_0, Nature (weakly) suffers but the firm (weakly) benefits.

In the relaxed model, the distribution of the demand during lead time satisfies the following property:

$$\Pr\left(\tilde{\xi} = (l+1-i)\,L + iH\right) = \frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}\,(1-\beta)^{i}\,.$$

The firm's time-average cost, which is denoted by $M\,(\beta, s_l)$, equals

$$M\,(\beta, s_l) = \sum_{i=0}^{l+1} \frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}\,(1-\beta)^{i} \begin{bmatrix} h\,(s_l - (l+1-i)\,L - iH)^{+} \\ +b\,((l+1-i)\,L + iH - s_l)^{+} \end{bmatrix}.$$

The firm chooses her base-stock level s_l to minimize $M\,(\beta, s_l)$ while Nature chooses β (since each two-point distribution in Ω_0 is uniquely determined by the parameter β) to maximize $M\,(\beta, s_l)$. For $l \geq 0$, let

$$\beta^{*} = \left(\frac{b}{b+h}\right)^{\frac{1}{l+1}}$$

be a constant, where the subscript l indicates the length of the lead time.

Proposition 1 *In the firm's relaxed model with backorders, the equilibrium strategy of Nature is the following distribution:*

$$F_l^{*} = \begin{cases} \Pr\left(\tilde{\theta} = \mu - \sigma\sqrt{\frac{1-\beta^{*}}{\beta^{*}}} \overset{def}{=} L_l\right) = \beta^{*}, \\ \Pr\left(\tilde{\theta} = \mu + \sigma\sqrt{\frac{\beta^{*}}{1-\beta^{*}}} \overset{def}{=} H_l\right) = 1-\beta^{*}. \end{cases} \tag{2.5}$$

The firm's equilibrium strategy is the following base-stock level:

$$s_l^{*} = (l+1)\mu + \sigma\left(\frac{2\beta^{*}-1}{2\sqrt{(1-\beta^{*})\beta^{*}}} - l\sqrt{\frac{1-\beta^{*}}{\beta^{*}}}\right). \tag{2.6}$$

The value of the zero-sum game is equal to

$$Z^{*} = b\sigma(l+1)\sqrt{\frac{1-\beta^{*}}{\beta^{*}}}. \tag{2.7}$$

Proposition 1 characterizes the equilibrium when Nature chooses a demand distribution from Ω_0. First, we find that the probability of a low demand (measured by β^{*}) is concave and increasing with respect to the lead time l. If l approaches infinity, β^{*} approaches 1. We can verify that the skewness of this distribution equals $\frac{2\beta^{*}-1}{\sqrt{\beta^{*}(1-\beta^{*})}}$, which approaches infinity in an asymptotic case where the lead time is sufficiently long. This skewed demand distribution is detrimental to the firm and forces the firm to raise her base-stock level.

Second, Proposition 1 suggests that when each $\tilde{\theta}_i$ is an iid two-point distribution, the demand during the lead time $\sum_{i=1}^{l+1} \tilde{\theta}_i$ is a transformed binomial distribution satisfying $\Pr\left(\tilde{\theta}_1 + \tilde{\theta}_2 + \ldots + \tilde{\theta}_{l+1} = (l+1)\, L_l\right) = \frac{b}{b+h}$, which is the firm's newsvendor ratio. This implies that in the steady-state equilibrium, backorders occur with probability $\frac{h}{b+h}$. Third, the value of the zero-sum game (i.e., the firm's optimized expected cost) is increasing with respect to the standard deviation σ, backorder cost b, and holding cost h but is unaffected by the mean μ.

Remark 1 *A crucial question is whether the relaxed solution indeed characterizes the true equilibrium. The key step is to investigate whether Nature changes its strategy when its action space expands from Ω_0 to Ω. If Nature does not alter its strategy, then the relaxed solution describes the true equilibrium. In the online appendix, we show that with backorders, Nature indeed maintains its strategy unchanged.*

3. Case With Lost Sales and Constant Lead Time

In the lost-sales model, the selling price of the firm's product is \$$p$, the production cost is \$$c$ (where $p > c \geq 0$), and the holding cost is \$$h > 0$. Thus, $(p - c)$ can be regarded as the penalty cost of lost sales. The sequence of the event is the same as that in the backorder model except that the demand in excess of the available on-hand inventory is lost rather than backordered. To avoid an uninteresting equilibrium, we assume that

$$\frac{p-c}{h} \geq \max\left\{\rho^2, l\right\}. \tag{3.1}$$

We justify condition (3.1) as follows. Both lost-sales and backorder models exclude the holding cost for the inventory in transit. However, if the holding cost for the inventory in transit is also \$$h$ per unit per period, the inequality in (3.1) implies that the profit margin $(p - c)$ must be sufficiently high to justify prestocking inventories. The second condition $(p - c) \geq \rho^2 h$ is consistent with Scarf's rule with zero lead time. Overall, condition (3.1) is a mild assumption on the cost parameters.

Given an unknown demand distribution, the firm is unable to estimate the transition probability under any state-dependent policy. Therefore, we focus on stationary base-stock policies such that the sum of the pipeline inventory and the on-hand inventory equals a constant s, which represents the firm's base-stock level. We also consider a profit maximization problem with lost sales rather than a cost minimization problem because of the convenience in illustrating the analysis. Similar to the backorder model, we can formulate the zero-sum game under lost sales as: $Z = \max_{s \geq 0} \inf_{F \in \Omega} \{M(s|F)\}$, where $M(s|F)$ is the firm's time-average profit. However, the challenge is that we must first derive $M(s|F)$ explicitly.

3.1 System Dynamics

We introduce the following notations. At the beginning of period t, we use I_t to represent the on-hand inventory before receiving any due delivery and vector $\mathbf{Q}_t = (Q_t, Q_{t-1}, \ldots, Q_{t-l})$, which is a row vector with $l+1$ elements, to describe the pipeline inventory. Counting from left to right, the first variable Q_t represents the order that is placed in period t, the second variable Q_{t-1} represents the order that was placed in period $t-1$, and the final variable Q_{t-l} represents the order that was placed in period $t-l$ but is due to arrive in the current period t. At any period t, we can fully characterize the state of the system using $(\mathbf{Q}_t|I_t)$, exploiting the fact that the demand per time period is either L or H. This considerably simplifies the state transition matrix in the lost sales model. The state transition proceeds as follows. After the order Q_{t-l} arrives, the on-hand inventory in period t becomes $I_t + Q_{t-l}$. The sales quantity in period t equals $W_t = \min(\theta_t, I_t + Q_{t-l})$. In the steady state under a stationary base-stock policy, the sum of the on-hand inventory and pipeline inventory must be lower than the base-stock level s and thus, in period t the firm's order quantity equals $Q_t = s - I_t - \sum_{i=t-l}^{t-1} Q_i$. The closing inventory of period t (after the demand θ_t arrives) becomes the initial inventory of period $t+1$ and equals $I_{t+1} = (I_t + Q_{t-l} - \theta_t)^+$. The order to be placed in the next period $t+1$ equals

$$Q_{t+1} = s - I_{t+1} - \sum_{i=t-l+1}^{t} Q_i = s - (I_t + Q_{t-l} - \theta_t)^+ - \sum_{i=t-l+1}^{t} Q_i,$$

which depends on the realized demand θ_t in the current period t. When $\theta_t = H$, $(I_t + Q_{t-l} - \theta_t)^+ = 0$, and hence $Q_{t+1} = s - \sum_{i=t-l+1}^{t} Q_i$. This expression is either $s - lL$, when all the other Q_i is L, or L when one of the Q_i is $s - lL$ and the rest are L.

Let $E(I_\infty)$, $E(W_\infty)$, and $E(Q_\infty)$ respectively represent the expected on-hand inventory, sales quantity, and order quantity in the steady state (where the subscript ∞ indicates that time approaches infinity). It is well known that $E(W_\infty) = E(Q_\infty)$ such that the inventory in the steady state never indefinitely builds up or depletes. The time-average profit then equals $M(F,s) = (p-c)E(W_\infty) - hE(I_\infty)$. The extant literature recognizes that the expressions $E(I_\infty)$ and $E(W_\infty)$ can be so complex as to be intractable when the iid demand follows an arbitrary distribution $F \in \Omega$. Therefore, we continue to focus on the relaxed model where Nature is restricted to choose a two-point distribution from set Ω_0.

3.2 Lost-Sales Equilibrium

The key step is to compute the time-average profit $M(\beta,s)$ (we use $M(\beta,s)$ to replace $M(s|F)$ since β uniquely determines Nature's two-point distribution). To determine the recurring states in the Markov chain, we start with a special case where $l = 0$. Evidently, $L \leq s \leq H$ must hold; otherwise, the firm can increase or decrease the base-stock level s to increase her profit. Thus, two recurring states occur: 1) with probability $(1-\beta)$, $I_t = (s-H)^+ = 0$; hence, $Q_t = s$; 2) with probability β, $I_t = s - L$; hence, $Q_t = L$. Two recurring states exist in this Markov chain, where state 1 is $(s|0)$ and state 2 is $(L|s-L)$. The steady-state distribution is $\pi_1 = 1-\beta$ and $\pi_2 = \beta$.

Next, we consider $l \geq 1$. Notably, when Nature plays a two-point distribution with a low realized demand L, the demand during the lead time must be at least $(l+1)L$ units. Because the max-min decision rule is conservative, we first consider a low inventory level with $(l+1)L \leq s \leq lL + H$ and verify this condition retrospectively. For the case where $l = 0$, this condition nonetheless holds. We find that the Markov chain now has $l+2$ recurring states, including $(s-lL, L, ..., L|0)$, $(L, s-lL, L, ..., L|0)$, ..., $(L, ..., L, s-lL|0)$, and $(L, ..., L|s-(l+1)L)$. Specifically, the initial inventory at the beginning of period t is either $I_t = 0$ (indicating that the previous period stockouts) or $I_t = s-(l+1)L$ (due to the precondition $s \leq lL + H$). When $I_t = 0$, only one of the orders in the pipeline can be $s - lL$ while all the other orders in the pipeline equal L such that $Q_t = L$. We recall that $\mathbf{Q}_t$ is a row vector with $(l+1)$ elements. Depending on which pipeline order equals $(s-lL)$, we have $(l+1)$ different $\mathbf{Q}_t$ vectors when $I_t = 0$. Similarly, if the on-hand inventory is $I_t = s-(l+1)L$, all the pipeline orders equal L including $Q_t = L$ and thus, $\mathbf{Q}_t =$

$(L, L, ..., L)$ is unique. In summary, the $l+2$ recurring states include $(s-lL, L, ..., L|0)$, referred to as state 1, $(L, s-lL, L, ..., L|0)$, referred to as state 2, ..., and finally $(L, ..., L|s-(l+1)L)$, referred to as state $l+2$.

Let matrix $\mathbf{I}_l$ be an identity matrix with l rows and l columns, vector $\mathbf{0}_l = (0, 0, \cdots, 0)$ be the zero row vector with l columns, and vector $\mathbf{0}_l^T$ be the zero column vector with l rows. The transition matrix is the following:

$$\begin{pmatrix} \mathbf{0}_l^T & \mathbf{I}_l & \mathbf{0}_l^T \\ 1-\beta & \mathbf{0}_l & \beta \\ 1-\beta & \mathbf{0}_l & \beta \end{pmatrix}. \tag{3.2}$$

For the first $l+1$ states, since only one of the pipeline orders can be $s-lL$, the state evolves from i to $(i+1)$, for $i = 1, ..., l$, with probability one, therefore the transition is represented by an identity matrix $\mathbf{I}_l$. For the transition from states $(l+1)$ and $(l+2)$, We elaborate the transition matrix using the last state $(L, ..., L|s-(l+1)L)$ as an example. Observe that $Q_{t-1} = \ldots = Q_{t-l} = L$, and $I_t = s-(l+1)L$. According to the base-stock policy, the order to be placed at time t satisfies

$$Q_t = s - I_t - \sum_{i=t-l}^{t-1} Q_i = s - s + (l+1)L - l \cdot L = L,$$

confirming that $Q_t = L$. After receiving the delivery Q_{t-l}, which is due at time t, the on-hand inventory becomes $I_t + Q_{t-l} = s-(l+1)L + L = s - lL$, which is positive. With probability $1-\beta$, the realized demand at period t is $\theta_t = H$. Because the firm stockouts at the end of period t, the leftover inventory is $I_{t+1} = 0$. At period $t+1$, the sum of the pipeline inventory and on-hand inventory before an order is placed equals $I_{t+1} + \sum_{i=t-l-1}^{t} Q_i = lL$. As such, the firm orders $Q_{t+1} = s - lL$ units of inventory in period $t+1$, and the Markov chain enters the state $(s-lL, L, ..., L|0)$. With probability β, the realized demand at time t is $\theta_t = L$; consequently, the leftover inventory is $I_{t+1} = s-(l+1)L > 0$. At time $t+1$, the firm orders $Q_{t+1} = L$ units of inventory such that the Markov chain returns to the same state $(L, ..., L|s-(l+1)L)$. Similarly, the transitions of the remaining states can be confirmed.

Using the Markov transition matrix in equation (3.2), we find that the steady-state distribution satisfies

$$\pi_1 = \pi_2 = \ldots = \pi_{l+1} = \frac{1}{(l+1) + \frac{\beta}{1-\beta}} = \frac{1-\beta}{l+1-l\beta}, \; \pi_{l+2} = \frac{\frac{\beta}{1-\beta}}{(l+1) + \frac{\beta}{1-\beta}} = \frac{\beta}{l+1-l\beta}. \tag{3.3}$$

In the notable special case where $l = 0$, equation (3.3) reduces to

$$\pi_1 = \frac{1}{1 + \frac{\beta}{1-\beta}} = 1 - \beta \text{ and } \pi_2 = \frac{\frac{\beta}{1-\beta}}{1 + \frac{\beta}{1-\beta}} = \beta,$$

confirming with our earlier discussion. The steady-state distribution in equation (3.3) enables us to precisely compute the time-average profit as follows (we refer readers to the proof of Proposition 2 for details):

$$M(\beta, s) = \frac{1}{l + 1 - l\beta} \left[(p - c + (l+1)h)\beta L - \beta(p - c + h)s + (p - c)s \right]. \tag{3.4}$$

We characterize the equilibrium in the following proposition.

Proposition 2 *If condition (3.1) holds, then the equilibrium in the firm's relaxed model with lost sales exhibits the following properties. i) The equilibrium strategy of Nature in each period is the following two-point distribution*

$$F^* = \begin{cases} \Pr\left(\tilde{\theta} = \mu - \sigma\sqrt{\frac{h}{p-c}} = L^*\right) = \frac{p-c}{p-c+h}, \\ \Pr\left(\tilde{\theta} = \mu + \sigma\sqrt{\frac{p-c}{h}} = H^*\right) = \frac{h}{p-c+h}. \end{cases}, \tag{3.5}$$

which is unaffected by the lead time. ii) The value of the zero-sum game equals $Z^* = (p-c)\mu - \sigma\sqrt{(p-c)h}$*, which is also unaffected by the lead time. iii) The equilibrium strategy of the firm is given by the following base-stock level:*

$$s^* = (l+1)\mu + \sigma\left(\frac{1}{2}\sqrt{\frac{p-c}{h}} - \frac{l+1}{2}\sqrt{\frac{h}{p-c}}\right). \tag{3.6}$$

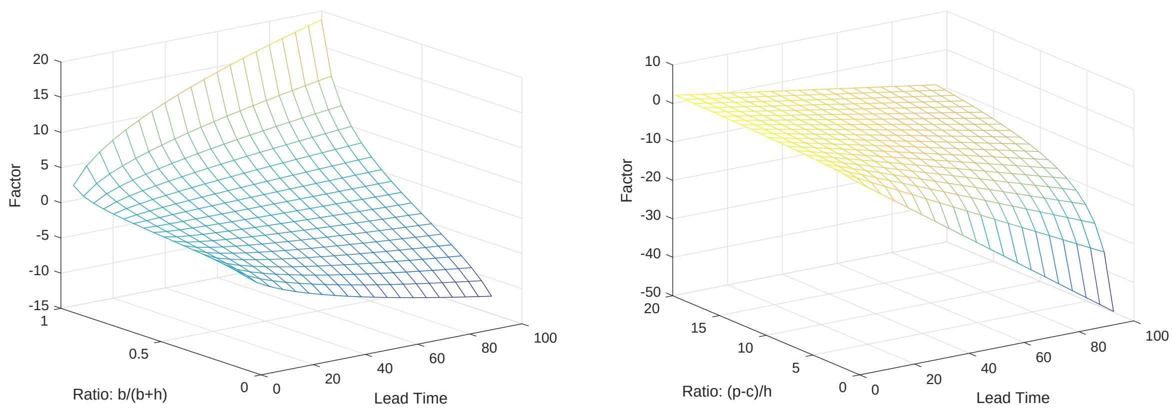
3.3 Backorders vs. Lost Sales

We now compare the equilibriums in Propositions 1 and 2. Both equations (2.6) and (3.6) are applicable for any lead time $l \geq 0$ and simple enough for Excel spreadsheet implementation. We define the safety stock factors as follows:

$$SF_b = \frac{2\beta^* - 1}{2\sqrt{(1-\beta^*)\beta^*}} - l\sqrt{\frac{1-\beta^*}{\beta^*}} \text{ and } SF_l = \frac{1}{2}\sqrt{\frac{p-c}{h}} - \frac{l+1}{2}\sqrt{\frac{h}{p-c}},$$

where the subscript b indicates backorders and l indicates lost sales. Figure 1 contrasts the safety stock factors when the lead time and newsvendor ratio vary. Figure 1a) shows that under backorders, the safety

Figure 1 Safety Stock Factors



(a) Backorders (b) Lost Sales

stock factor SF_b is nonmonotonic with respect to the lead time or newsvendor ratio; however, Figure 1b) shows that under lost sales, the safety stock factor SF_l is nonincreasing with respect to the lead time.

More important, Nature behaves differently. Lost sales force Nature to resume the game whenever stock-out occurs. If the distribution F^* in equation (3.5) is the most unfavorable for the firm with zero lead time, Nature continues to play the same distribution F^* in each period despite the longer lead time, explaining why the value of the zero-sum game is unaffected by the lead time due to lost sales. In contrast, backorders create an accumulative effect, which exacerbates the shortage problem and enables Nature to adjust the parameter β in its two-point distribution to punish the firm when the lead time l increases. Therefore, backorders are more detrimental to the firm than lost sales.

Corollary 1 *In the relaxed models, if the marginal understock and overstock costs remain unchanged, backorders are more costly than lost sales from the worst-case perspective.*

3.4 Constant Order Policy (COP)

A surprising result that a COP can be asymptotically optimal with a long lead time has inspired a growing literature stream. Because the optimal COP quantity is difficult to determine, Bu et al. (2020) proposed the following formula:

$$R = \mu\left(1 - \sqrt{\frac{\rho^2}{1+\frac{2c_u}{c_o}}}\right) = \mu\left(1 - \sqrt{\frac{h\rho^2}{h+2(p-c)}}\right), \tag{3.7}$$

where R is the constant order quantity, which is *unaffected* by the lead time. An important result in Proposition 2 is that Nature's equilibrium strategy does not change with the lead time. The two-point distribution played by Nature offers a helpful hint on choosing the COP quantity. We propose a new closed-form formula $R' = \mu\left(1-\rho\sqrt{\frac{h}{p-c}}\right)$, which is the lower realized demand that Nature chooses. Interestingly, our numerical experiments indicate that the new COP quantity R' substantially outperforms R when the lead time is short.

4. Numerical Study

Because the backorder model is closely related to the risk-pooling model developed in Section 5.1, where we present the numerical study about equation (2.6), this section only considers the lost-sales model.

4.1 Lost-Sales Model

We evaluate the performance of the robust base-stock policy shown in equation (3.6). We assume that the production cost is $c = 1$ and the holding cost is $h = 1$. We consider four different distributions for the underline distribution F, including the Poisson, triangle, uniform, and exponential distributions. All of these distributions have the same mean $\mu = 5$. The Poisson distribution is discrete and has a coefficient of variation $\rho = 0.447$. The remaining three distributions are nonnegative and continuous. We do not use a normal distribution because it could be negative. The coefficient of variation of the exponential distribution is $\rho = 1$, that of the chosen triangle distribution is $\rho = 0.408$, and that of the uniform distribution is $\rho = 0.577$. We compute the robust base-stock level s^* according to equation (3.6). Because the lost-sales model in an unambiguous environment is notoriously difficult, we use the sample path method to evaluate the performance of the robust policy. We generate four sample paths according to the underline distribution shown in the first column of Table 1. Each sample path consists of 400 periods. We disclose the sample paths that we use in the supplementary file (DataDisclosure.xlsx) so that interested readers can independently replicate our results. To avoid the technical challenge in computing the optimal profit under a known distribution, we use the retrospective optimization method to compute $\hat{Z}$ as a benchmark. Specifically, we choose a static (rather than dynamic) base-stock level $\bar{s}$ based on the generated sample path to maximize the average profit of the entire path. Clearly, the value of $\hat{Z}$ must be larger than the optimal profit under the best base-stock policy because the retrospective method chooses the static base-stock level from hindsight. The performance gap is defined as $Gap = \frac{\hat{Z} - Z\left(s^{*}, F\right)}{\hat{Z}} \times 100\%$, where $Z\left(s^{*}, F\right)$ is the value of the zero-sum game with lost sales under a specific distribution F. We report the data of Gap in Table 1.

Table 1 Performance Gap of Robust Policies with Lost Sales: We change the selling price to $p = 5$, 10, 20, and 30 so that the range of the newsvendor ratio is between 0.8 and 0.967. We also change the lead time to $l = 1$, 2, 3, and 4.

Distribution F	Lead Time $l = 1$				Lead Time $l = 2$			
	$p = 5$	$p = 10$	$p = 20$	$p = 30$	$p = 5$	$p = 10$	$p = 20$	$p = 30$
Poisson	0.9%	3.2%	1.0%	1.1%	0.2%	0.5%	0.5%	0.3%
Exponential	0.4%	0.0%	0.1%	0.0%	0.2%	0.5%	0.4%	0.0%
Triangle	1.8%	2.5%	0.7%	1.4%	0.8%	0.5%	0.1%	0.0%
Uniform	0.3%	2.5%	2.8%	2.0%	1.8%	2.5%	1.3%	0.5%
	Lead Time $l = 3$				Lead Time $l = 4$			
Poisson	0.0%	1.2%	0.5%	0.1%	0.1%	0.7%	0.8%	0.5%
Exponential	1.0%	0.3%	0.4%	0.6%	0.3%	1.8%	2.7%	0.2%
Triangle	0.3%	0.3%	0.3%	0.2%	0.1%	0.2%	0.6%	0.4%
Uniform	1.2%	2.9%	1.8%	0.9%	0.7%	2.7%	2.3%	1.5%

When the demand is exponential, Table 1 shows that the average Gap is 0.6%, meaning that the average profit under the robust base-stock policy s^{*} is on average 0.6% lower than the retrospective optimal profit $\hat{Z}$. When the demand follows the Poisson distribution, Table 1 shows that the average Gap is 0.7%. A notable phenomenon is that the robust policy performs the worst under the uniform distribution (with an average gap of 1.74%), suggesting that the uniform distribution could be the most rigorous benchmark for testing the robust policy.

4.2 Industrial Data and Heuristics

The extant literature has proposed various heuristic methods with different levels of complexity to solve the lost-sales model. We choose the following two representative methods based on their performance and implementation simplicity. 1) The weighted-average policy (Huh et al. 2009), under which the stationary base-stock level is equal to

$$s_w = \frac{p-c}{p-c+h} F_{l+1}^{-1}\left(\frac{p-c}{p-c+h}\right) + \frac{h}{p-c+h} F^{-1}\left(\frac{p-c}{p-c+h}\right), \tag{4.1}$$

where F^{-1} is the inverse of the cumulative distribution function and F_{l+1}^{-1} is the inverse of the $(l+1)$ fold convolution. Notably, in equation (4.1), two *backorder* solutions are applied to construct a weighted-average base-stock level for the *lost-sales* counterpart. The first backorder solution assumes the lead time to be l while the second backorder solution assumes a zero lead time. The weight to be put on the first backorder solution is the newsvendor ratio. 2) The COP quantities R and R' shown in Section 3.4. In terms of computational effort, s_w, R and R' are light and can be implemented in Excel (but s_w requires a prior demand distribution and its $(l+1)$-fold convolution). When the demand distribution is unknown, many sophisticated state-dependent policies such as the projected inventory level policy in van Jaarsveld and Arts (2021) cannot be easily implemented.

We use daily data from a large beverage company in China (`https://www.coap.online/competitions/1`) to conduct an applied numerical study. The data set contains the demand data for seventy-seven SKUs (stock-keeping units). We extract the data for SKU8 for testing purposes for the following reasons. First, SKU8 has a long time frame (starting from 1 January 2018 to 31 July 2020) to ensure the steady state can be reached. Second, for SKU8, there are 142 zero-demand days (significantly lower than many other SKUs). During the lockdown period related to COVID-19 restrictions in China, days with zero demand are common but can undermine the performance of the other candidate policies. For example, inventory accumulates when the demand is zero under the COP policy. We find that the sample mean is $1,824$ and sample standard deviation is $1,464$. We draw a histogram to confirm that the Poisson distribution is a suitable prior distribution so that we can compute s_w according to equation (4.1). Readers can refer to the online supplementary file (DataDisclosure.xlsx) for details.

The original data exclude pricing information due to business confidentiality. We assume that $p = 5$, $c = 1$ and $h = 1$ so that the newsvendor ratio is 0.8. To benchmark the performance of various methods, we use the retrospective optimization method to determine a stationary base-stock level s_r and compute the retrospective profit $\hat{Z} = Z(s_r)$ using the industrial data. When the demand distribution is unknown, the retrospective profit $\hat{Z}$ provides a convenient benchmark. Due to the positive lead time, we assume that the firm starts with zero pipeline and on-hand inventory. However, non-zero initial inventories do not qualitatively affect the ranking of the performance due to the long study period. The following Table 2 reports the values of $Z(s^*)$, $Z(s_w)$, $Z(R)$ and $Z(R')$ when the lead time changes from $l = 1$ to $l = 4$.

Table 2 Performance Gaps in the Industrial Data

Lead	Retrospective		Our Solution		Weighted Average		Constant Order Policy			
Time	s_r	$\hat{Z}$	s^*	$Z(s^*)$	s_w	$Z(s_w)$	R	$Z(R)$	R'	$Z(R')$
$l = 1$	4412	4309	4381	4308	3331	4015	1336	2519	1092	3551
$l = 2$	5764	3924	5839	3923	4799	3798	1336	2512	1092	3545
$l = 3$	6734	3457	7296	3427	6266	3426	1336	2505	1092	3540
$l = 4$	8303	2989	8754	2983	7733	2971	1336	2498	1092	3534
Average Profit		3670		3660		3552		2508		3542

The last row of Table 2 shows that our robust policy delivers the best performance in terms of average profit across all lead times. We also examine the maximum percentage gap of each policy. We find that 1) the maximum gap of the robust policy s^* is 0.9% (when the lead time is $l = 3$); 2) that of the weighted-average policy s_w is 6.8% (when the lead time is $l = 1$); 3) that of the COP heuristic R is 41.5% (when the lead time is $l = 1$); and 4) that of the second COP heuristic R' is 17.6% (when the lead time is $l = 1$). In summary, given its consistent performance and light computational effort, we recommend equation (3.6) for practical implementation.

An interesting observation is that the second COP policy R' delivers a retrospective profit higher than $\hat{Z}$ when the lead time is $l = 3$ or $l = 4$. We emphasize that the true optimal policy for the lost-sales model has a complex structure and is state-dependent (for example, see Zipkin 2008a). Theoretically, neither the COP nor the stationary base-stock policy dominates. The retrospective profit $\hat{Z}$ is based on a stationary base-stock policy, whereas $Z(R')$ is based on a constant order quantity policy. Hence, the difference $Z(R') - \hat{Z}$ can have an indefinite sign. Different inventory policies may produce different equilibria. We plan to investigate the equilibrium under the COP policy in the future.

5. Applications and Extensions

The multiperiod inventory system with a constant replenishment lead time is a fundamental model in the field of Operations Management. Furthermore, the zero-sum game method (using two-point iid distributions) appears to overcome the technical difficulties caused by a high-dimensional convolution associated with many interesting stochastic control problems. This section explores the application of this method to other classes of operational problems.

5.1 A Single-Period Risk-Pooling System

We consider a single-period system with a (central) warehouse and $N \geq 1$ retailers. Let θ_i be the demand of retailer i (where $i = 1, 2, ..., N$). We assume that θ_i is iid with mean μ and variance σ^2. Let $\xi = \sum_{i=1}^{N} \theta_i$ be the aggregate demand that the warehouse receives and s be the inventory level of the warehouse. We assume that each retailer is equally important. If a shortage occurs, the warehouse linearly allocates the available inventory according to θ_i. The aggregate shortage of the warehouse equals $(\xi - s)^+$ and the left-over inventory equals $(s - \xi)^+$. Let c_u be the understock cost and c_o be the overstock cost of the warehouse. By properly defining the values of c_u and c_o, we can comprehensively analyze either backorder or lost sales in this single-period model. Under the max-min decision rule, the warehouse determines the inventory level by solving the following model:

$$Z = \min_{s \geq 0} \sup_{F \in \Omega} \left\{ \int_0^\infty c_u (\xi - s)^+ + c_o (s - \xi)^+ \, dF_N(\xi) \right\}, \tag{5.1}$$

where F_N is the N-fold convolution of iid demand distribution F.

The literature has followed Scarf's rule by assuming that Nature directly chooses the distribution of ξ subject to the mean constraint $E(\tilde{\xi}) = N\mu$ and variance constraint $Var(\tilde{\xi}) = N\sigma^2$. As such, the inventory level and the value of the zero-sum game of the warehouse satisfy

$$\bar{s} = N\mu + \frac{\sqrt{N}\sigma}{2}\left(\sqrt{\frac{c_u}{c_o}} - \sqrt{\frac{c_o}{c_u}}\right) \text{ and } \bar{Z} = \sqrt{Nc_u c_o}\sigma. \tag{5.2}$$

Notably, this approach uses only the correlation information but ignores the fact that F_N is the convolution of N iid distributions, each of which may be highly skewed, although F_N is approximately symmetrical when N is large.

Alternatively, by letting $N = l + 1$, we can directly apply Proposition 1 to obtain the optimal inventory level

$$s_N^* = N\mu + \sigma\left(\frac{2\beta^* - 1}{2\sqrt{(1-\beta^*)\beta^*}} - (N-1)\sqrt{\frac{1-\beta^*}{\beta^*}}\right), \tag{5.3}$$

where $\beta^* = (c_u/(c_u + c_o))^{1/N}$ and the value of the zero-sum game equals $Z^* = c_u\sigma N\sqrt{\frac{1-\beta^*}{\beta^*}}$ under the iid constraints. From the zero-sum game perspective, the iid distribution constraints reduce Nature's action space and therefore benefit the firm. We can verify that $\bar{Z}$ is higher than Z^* for any ratio $\frac{c_u}{c_u+c_o}$.

5.1.1 Numerical Performance The values of $\bar{Z}$ and Z^* merely quantify the firm's costs under the worst distributions. We are more interested in the numerical performance of $\bar{s}$ and s_N^* in a known prior distribution. We set the holding cost at $h = 1$ and consider three commonly used distributions: (1) the normal distribution (which is continuous and symmetric); (2) the Poisson distribution (which is discrete and skewed); and (3) the exponential distribution (which is continuous and skewed). Both the normal and Poisson distributions have the same mean $\mu = 5$ and same standard deviation $\sigma = \sqrt{5}$. However, the mean and standard deviation of the exponential distribution are $\mu = \sigma = 1$. As such, the normal and Poisson distributions have the same coefficient of variation $\rho = 0.447$, and the exponential distribution has a coefficient of variation $\rho = 1$.

We compute $\bar{s}$ and s_N^* based on equations (5.2) to (5.3). We also compute the distribution-dependent base-stock levels s_n, s_p and s_e (where the subscript n indicates the normal distribution, p indicates the Poisson distribution, and e indicates the exponential distribution). We do not round $\bar{s}$ or s_N^* to an integer even when demand follows the Poisson distribution. Under the normal distribution, we compute $Z(\bar{s}|n)$, $Z(s_N^*|n)$, and $Z(s_n|n)$, which respectively represent the expected cost when the base-stock levels $\bar{s}$, s_N^* and s_n are implemented. We compute the performance gaps: $Gap\text{-}\bar{s} = \frac{Z(\bar{s}|n)-Z(s_n|n)}{Z(s_n|n)} \times 100\%$ and $Gap\text{-}s_N^* = \frac{Z(s_N^*|n)-Z(s_n|n)}{Z(s_n|n)} \times 100\%$. Similarly, under the Poisson and exponential distributions, we also compute the two performance gaps using the distribution-dependent profits $Z(s_p|p)$ and $Z(s_e|e)$ as the benchmarks. We report the results in Table 3.

Table 3 Performance of Robust Base-Stock Levels with Risk-Pooling: $h=1$

Backorder Cost	N	Normal Distribution			Poisson Distribution			Exponential Distribution		
		Gap-$\bar{s}$	Gap-s_N^*	$Z(s_n\|n)$	Gap-$\bar{s}$	Gap-s_N^*	$Z(s_p\|p)$	Gap-$\bar{s}$	Gap-s_N^*	$Z(s_e\|e)$
$b=1$	2	0.0%	0.88%	2.52	0.0%	1.42%	2.50	2.9%	0.5%	1.05
	3	0.0%	1.42%	3.09	0.0%	1.47%	3.07	1.9%	0.0%	1.32
	4	0.0%	1.70%	3.57	0.0%	1.40%	3.55	1.4%	0.0%	1.54
	5	0.0%	1.88%	3.99	0.0%	1.31%	3.98	1.1%	0.1%	1.74
	6	0.0%	2.00%	4.37	0.0%	1.60%	4.36	0.9%	0.3%	1.91
$b=4$	2	0.43%	1.62%	4.43	0.58%	0.82%	4.61	0.06%	0.06%	2.24
	3	0.43%	2.15%	5.42	0.44%	2.43%	5.59	0.0%	0.39%	2.71
	4	0.43%	2.43%	6.26	0.63%	2.25%	6.44	0.0%	0.69%	3.10
	5	0.43%	2.61%	7.00	0.64%	2.34%	7.17	0.02%	0.94%	3.44
	6	0.43%	2.74%	7.67	0.15%	2.47%	7.86	0.03%	1.13%	3.75
$b=9$	2	0.13%	0.00%	5.55	0.61%	0.08%	5.87	0.0%	0.12%	3.09
	3	0.13%	0.04%	6.80	0.39%	0.49%	7.12	0.0%	0.23%	3.68
	4	0.13%	0.07%	7.85	0.06%	0.65%	8.19	0.0%	0.30%	4.17
	5	0.13%	0.09%	8.77	0.00%	0.01%	9.15	0.0%	0.34%	4.59
	6	0.13%	0.11%	9.61	0.33%	0.47%	9.95	0.0%	0.36%	4.97

Due to the relationship between the lead time and the number of iid retailers (i.e., $l = N - 1$), Table 3 also evaluates the performance of equation (2.6) in the backorder model. In a special case where $N = 1$ (such that the lead time is $l = 0$), $\bar{s} = s_N^*$ must hold and hence, we start with $N = 2$. Table 3 shows that neither $\bar{s}$ nor s_N^* dominates but both of them perform satisfactorily. In general, when the prior distribution is symmetric, $\bar{s}$ is more likely to outperform s_N^* (because s_N^* anticipates that Nature chooses a skewed distribution in the zero-sum game).

5.2 A Two-Echelon System with Expedited Delivery

We consider a two-echelon system with one central warehouse and N iid retailers. The (regular) lead time of the warehouse is l_w and that from the warehouse to the retailer k is l_k (where $k = 1, 2, ..., N$). Each retailer faces iid demand with mean μ_k, variance σ_k^2, backorder cost b_k and holding cost h_k. Let $\beta_k^* = (b_k/(b_k + h_k))^{1/(l_k+1)}$ be the constant. The warehouse and retailers independently decide their own stationary base-stock levels. Proposition 1 indicates that the base-stock level of retailer k equals:

$$s_{l_k}^* = (l_k + 1)\mu_k + \sigma_k \left(\frac{2\beta_k^* - 1}{2\sqrt{(1 - \beta_k^*)\beta_k^*}} - l\sqrt{\frac{1 - \beta_k^*}{\beta_k^*}} \right).$$

We assume the warehouse uses a stationary base-stock policy to replenish the orders from the retailers, with holding and ordering cost of h_w and c_w respectively. However, the shortfalls are filled by a backup vendor at a cost of $p_w > c_w$ for expedited deliveries (zero lead time) such that the warehouse fills all the orders arriving from the downstream retailers. Under the assumption that retailers' demands are independent of one another and every retailer employs a stationary base-stock policy, in the steady state, each retailer places an order that equals the realized demand in the previous period. We observe that this two-echelon system can be decomposed into two separate locations. The warehouse operates as if with lost sales, and hence we can use Proposition 2 to compute the base-stock level for the warehouse as follows:

$$s_w^* = (l_w + 1)\sum_k \mu_k + \sqrt{\sum_k \sigma_k^2} \left(\frac{1}{2}\sqrt{\frac{p_w - c_w}{h_w}} - \frac{l_w + 1}{2}\sqrt{\frac{h_w}{p_w - c_w}} \right).$$

Certainly, if expedited delivery is unavailable, the system cannot be decomposed and becomes a variant of the classic serial supply chain. The new challenge is that the demand distribution is endogenous in the zero-sum game, and we plan to study this system in future research.

6. Conclusion

This paper considers the classic inventory models with a positive lead time when only the mean and variance of the demand are known. We establish multiperiod zero-sum games to obtain the firm's distributionally robust base-stock policies in closed forms. Under backorders, we show that Nature's equilibrium strategy is a two-point distribution, which is affected by the lead time. We show that a myopic policy is optimal under backorders and the robust base-stock level is indeed optimal among all feasible policies.

Due to the notorious difficulty of the lost-sales model and the success in solving the backorder model, we consider a relaxed model where Nature chooses only two-point distributions under lost sales. We fully characterize the relaxed zero-sum game equilibrium to show that Nature's optimal two-point distribution is unaffected by the lead time. This result differs from the backorder model. Lost sales avoid the accumulative effect that backorders entail such that Nature cannot punish the firm by changing its distribution. Therefore, backorders exacerbate the shortage problem of the firm when the lead time increases, making backorders more costly than lost sales. Although the complete proof of the general model of lost sales is still unavailable, we believe that Nature's strategy remains to be a two-point distribution in the zero-sum game with lost sales. A tractable special case also confirms this conjecture. Nonetheless, the numerical study shows that our robust base-stock policy outperforms rival policies and can serve as a benchmark for future research.

This paper develops a promising methodology to analyze a large class of robust inventory models. Traditionally, the objective of the research on robust inventory models is to provide an easy-to-compute solution for the unambiguous model. This paper tackles a notoriously difficult problem, in which the unambiguous counterpart is intractable. Therefore, our closed-form solution can serve as the theoretical replacement of the unambiguous solution. Alternatively, we can ignore the prior distribution and use its mean and variance to compute the robust base-stock level.

References

Aizenman, M., Germinet, F., Klein, A., and Warzel, S. (2022). On bernoulli decompositions for random variables, concentration bounds, and spectral localization. *Working Paper*.

Arts, J., Levi, R., van Houtum, G.-J., and Zwart, B. (2015). Base-stock policies for lost-sales models: Aggregation and asymptotics. *BETA publicatie: working papers*, Vol. 491.

Bu, J., Gong, X., and Yao, D. (2020). Constant-order policies for lost-sales inventory models with random supply functions: Asymptotics and heuristic. *Operations Research*, 68(4):1063–1073.

Chen, W., Dawande, M., and Janakiraman, G. (2014). Fixed-dimensional stochastic dynamic programs: An approximation scheme and an inventory application. *Operations Research*, 62(1):81–103.

Corsten, D. and Gruen, T. W. (2004). Stock-outs cause walkouts. *Harvard Business Review*, 82(5):26–28.

Ergin, E., Gümüş, M., and Yang, N. (2022). An empirical analysis of intra-firm product substitutability in fashion retailing. *Production and Operations Management*, 31(2):607–621.

Fudenberg, D. and Tirole, J. (1991). *Game Theory*. The MIT Press, Cambridge, Massachusetts.

Gallego, G. and Moon, I. (1993). The distribution free newsboy problem: Review and extensions. *Journal of the Operational Research Society*, pages 825–834.

Goldberg, D. A., Katz-Rogozhnikov, D. A., Lu, Y., Sharma, M., and Squillante, M. S. (2016). Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. *Mathematics of Operations Research*, 41(3):898–913.

Goldberg, D. A., Reiman, M. I., and Wang, Q. (2021). A survey of recent progress in the asymptotic analysis of inventory systems. *Production and Operations Management*, 30(6):1718–1750.

Huh, W. T., Janakiraman, G., Muckstadt, J. A., and Rusmevichientong, P. (2009). Asymptotic optimality of order-up-to policies in lost sales inventory systems. *Management Science*, 55(3):404–420.

Janakiraman, G. and Roundy, R. O. (2004). Lost-sales problems with stochastic lead times: Convexity results for base-stock policies. *Operations Research*, 52(5):795–803.

Levi, R., Janakiraman, G., and Nagarajan, M. (2008). A 2-approximation algorithm for stochastic inventory control models with lost sales. *Mathematics of Operations Research*, 33(2):351–374.

Li, Z., Fu, Q., and Teo, C.-P. (2022). Closed-form solutions for distributionally robust inventory management: Extended reformulation using zero-sum game. *The University of Sydney Working Paper*, X(1):1–53.

Mamani, H., Nassiri, S., and Wagner, M. R. (2017). Closed-form solutions for robust inventory management. *Management Science*, 63(5):1625–1643.

Scarf, H. (1958). A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production*, 10(2):201–209.

Shang, K. H. and Song, J.-S. (2003). Newsvendor bounds and heuristic for optimal policies in serial supply chains. *Management Science*, 49(5):618–638.

Sobel, M. J. (1981). Myopic solutions of markov decision processes and stochastic games. *Operations Research*, 29(5):995–1009.

van Jaarsveld, W. and Arts, J. (2021). Projected inventory level policies for lost sales inventory systems: Asymptotic optimality in two regimes. *arXiv:2101.07519*.

Wei, L., Jasin, S., and Xin, L. (2021). On a deterministic approximation of inventory systems with sequential service-level constraints. *Operations Research*, 69(4):1057–1076.

Xin, L. (2021). Technical note–understanding the performance of capped base-stock policies in lost-sales inventory models. *Operations Research*, 69(1):61–70.

Xin, L. and Goldberg, D. A. (2016). Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Operations Research*, 64(6):1556–1565.

Xin, L. and Goldberg, D. A. (2018). Asymptotic optimality of tailored base-surge policies in dual-sourcing inventory systems. *Management Science*, 64(1):437–452.

Xin, L. and Goldberg, D. A. (2022). Distributionally robust inventory control when demand is a martingale. *Mathematics of Operations Research*, 47(3):2387–2414.

Zipkin, P. (2008a). Old and new methods for lost-sales inventory systems. *Operations Research*, 56(5):1256–1263.

Zipkin, P. (2008b). On the structure of lost-sales inventory models. *Operations research*, 56(4):937–944.

Appendix: Technical Proofs

Proof of Proposition 1:

We first derive the $(l+1)$-fold convolution of the two-point distribution specified by Ω_0 in (2.4). For expositional simplicity, let $\Delta = H - L$ and $\tilde{\tau}$ be the standard Bernoulli random variable with parameter $(1-\beta)$

such that $\Pr(\tilde{\tau}=1)=1-\beta$ and $\Pr(\tilde{\tau}=0)=\beta$. The convoluted variable $\tilde{\xi}$ satisfies $\xi=(l+1)L+\Delta\left(\sum_{i=0}^{l+1}\tilde{\tau}_i\right)$, which is a transformed binomial random variable. In particular,

$$\Pr\left(\tilde{\xi}=(l+1-i)L+iH\right)=\frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}(1-\beta)^i.$$

Therefore, the firm's time-average cost under backorder, which is denoted by $M(\beta,s)$, equals

$$M(\beta,s)=\sum_{i=0}^{l+1}\frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}(1-\beta)^i\begin{bmatrix}h(s-(l+1-i)L-iH)^+\\+b((l+1-i)L+iH-s)^+\end{bmatrix}.$$

The firm chooses her base-stock level s to minimize $M(\beta,s)$ while Nature chooses β (since each two-point distribution in Ω_0 is uniquely determined by the parameter β) to maximize $M(\beta,s)$. The two first order conditions (FOCs) with respect to β and s (i.e., $\frac{\partial M(\beta,s)}{\partial\beta}=\frac{\partial M(\beta,s)}{\partial s}=0$) are sufficient and necessary to characterize the equilibrium.

Before calculating the FOCs, we need to simplify $M(\beta,s)$. We start with the case with low inventory level by assuming that the base-stock level satisfies $(l+1)L<s\leq lL+H$. The transformed binomial distribution yields the following:

$$\begin{aligned}M(\beta,s)&=\beta^{l+1}h(s-(l+1)L)+\sum_{i=1}^{l+1}\frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}(1-\beta)^ib((l+1-i)L+iH-s)\\&=\beta^{l+1}h(s-(l+1)L)-\beta^{l+1}b((l+1)L-s)+\beta^{l+1}b((l+1)L-s)\\&+\sum_{i=1}^{l+1}\frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}(1-\beta)^ib((l+1-i)L+iH-s)\\&=\beta^{l+1}h(s-(l+1)L)-\beta^{l+1}b((l+1)L-s)\\&\quad+\sum_{i=0}^{l+1}\frac{(l+1)!}{(l+1-i)!i!}\beta^{l+1-i}(1-\beta)^ib((l+1-i)L+iH-s)\\&=\beta^{l+1}(h+b)(s-(l+1)L)+b((l+1)\mu-s)\\&=\beta^{l+1}(h+b)\left(s-(l+1)\mu+(l+1)\sigma\sqrt{\frac{1-\beta}{\beta}}\right)+b((l+1)\mu-s)\\&=(h+b)\left[\beta^{l+1}(s-(l+1)\mu)+\beta^l(l+1)\sigma\sqrt{(1-\beta)\beta}\right]+b((l+1)\mu-s),\end{aligned}$$

where the fourth equation is due to $E\left(\tilde{\xi}\right)=(l+1)\mu$. By letting $\frac{\partial M(\beta,s)}{\partial s}=(h+b)\beta^{l+1}-b=0$, we obtain $\beta^*=\left(\frac{b}{b+h}\right)^{\frac{1}{l+1}}$ to confirm equation (2.5).

The FOC with respect to β yields

$$\frac{\partial M(\beta,s)}{\partial\beta} = (h+b)\left[(l+1)\beta^l(s-(l+1)\mu) + \frac{(l+1)\sigma\beta^l(1-2\beta)}{2\sqrt{(1-\beta)\beta}} + l(l+1)\beta^{l-1}\sigma\sqrt{(1-\beta)\beta}\right] = 0.$$

The constant $(h+b)(l+1)$ can be factored out. We find that

$$\beta^l(s-(l+1)\mu) + \frac{\sigma\beta^l(1-2\beta)}{2\sqrt{(1-\beta)\beta}} + l\beta^{l-1}\sigma\sqrt{(1-\beta)\beta} = 0,$$

implying that

$$\beta^l(s-(l+1)\mu)2\sqrt{(1-\beta)\beta} + \sigma\beta^l(1-2\beta) + 2l\beta^{l-1}\sigma(1-\beta)\beta = 0,$$

which is equivalent to

$$(s-(l+1)\mu)2\sqrt{(1-\beta)\beta} + \sigma(1-2\beta) + 2l\sigma(1-\beta) = 0.$$

After re-organizing the terms, we confirm equation (2.6). Substituting β^* and s_l^* into $M(\beta,s)$, we find that the value of the zero-sum game equals $Z^* = b\sigma(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}}$.

Next, we need to show that L_l is nonnegative. Notice that $\frac{\beta^*}{1-\beta^*} = \frac{1}{1-(\frac{b}{b+h})^{\frac{1}{l+1}}} - 1$, which is increasing in l; therefore $\frac{\beta^*}{1-\beta^*} \geq \frac{\frac{b}{b+h}}{1-\frac{b}{b+h}} = \frac{b}{h}$. To prove $L_l = \mu - \sigma\sqrt{\frac{1-\beta^*}{\beta^*}} \geq 0$, it suffices to prove that $\frac{\beta^*}{1-\beta^*} \geq (\frac{\sigma}{\mu})^2 = \rho^2$. According to the assumption of $b \geq \rho^2 h$, we can conclude that $\frac{\beta^*}{1-\beta^*} \geq \frac{b}{h} \geq \rho^2$ for any $l \geq 0$.

The final step is to verify the presumption $(l+1)L_l < s_l^* \leq lL_l + H_l$. We have

$$\begin{aligned} s_l^* - (l+1)L_l &= (l+1)\mu + \sigma\left(\frac{2\beta^*-1}{2\sqrt{(1-\beta^*)\beta^*}} - l\sqrt{\frac{1-\beta^*}{\beta^*}}\right) - (l+1)\left(\mu - \sigma\sqrt{\frac{1-\beta^*}{\beta^*}}\right) \\ &= \sigma\left(\frac{2\beta^*-1}{2\sqrt{(1-\beta^*)\beta^*}} + \sqrt{\frac{1-\beta^*}{\beta^*}}\right) = \frac{1}{2\sqrt{(1-\beta^*)\beta^*}} > 0, \end{aligned}$$

and

$$\begin{aligned} lL_l + H_l - s_l^* &= (l+1)\mu + \sigma\left(\sqrt{\frac{\beta^*}{1-\beta^*}} - l\sqrt{\frac{1-\beta^*}{\beta^*}}\right) - (l+1)\mu - \sigma\left(\frac{2\beta^*-1}{2\sqrt{(1-\beta^*)\beta^*}} + l\sqrt{\frac{1-\beta^*}{\beta^*}}\right) \\ &= \sigma\left(\sqrt{\frac{\beta^*}{1-\beta^*}} - \frac{2\beta^*-1}{2\sqrt{(1-\beta^*)\beta^*}}\right) = \frac{1}{2\sqrt{(1-\beta^*)\beta^*}} > 0. \end{aligned}$$

Hence, equation (2.6) is the firm's equilibrium strategy. Q.E.D.

Uniqueness of Equilibrium in Proposition 1

We show that $M(\beta, s)$ is supermodular by verifying that $\frac{\partial^2 M}{\partial\beta\partial s} > 0$. With $l \geq 0$, we find that

$$\begin{aligned}\frac{\partial^2 M}{\partial\beta\partial s} &= \frac{\partial M}{\partial\beta}\left(\frac{\partial M}{\partial s}\right) = \frac{\partial}{\partial\beta}\left(c_o \Pr(\xi \leq s) - c_u \Pr(\xi > s)\right) \\ &= \frac{\partial}{\partial\beta}\left(c_o \Pr(\xi \leq s) - c_u + c_u \Pr(\xi \leq s)\right) = (c_o + c_h)\frac{\partial}{\partial\beta}\Pr(\xi \leq s).\end{aligned}$$

Here, $\frac{\partial M}{\partial s} = c_o \Pr\left(\tilde{\xi} \leq s\right) - c_u \Pr\left(\tilde{\xi} > s\right)$ is the well-known result in the newsvendor model where the random demand is $\tilde{\xi}$ and inventory level is s. There are $l+1$ cases with recursive structures.

Case 1) When $(l+1)L \leq s < lL + H$, $\Pr(\xi \leq s) = \beta^{l+1}$ and hence, $\frac{\partial}{\partial\beta}\Pr(\xi \leq s) = (l+1)\beta^l > 0$.

Case 2) When $lL + H < s < (l-1)L + 2H$, $\Pr(\xi \leq s) = \beta^{l+1} + (l+1)\beta^l(1-\beta) = (l+1)\beta^l - l\beta^{l+1}$ and hence,

$$\frac{\partial}{\partial\beta}\Pr(\xi \leq s) = (l+1)l\beta^{l-1} - (l+1)l\beta^l = (l+1)l\beta^{l-1}(1-\beta) > 0.$$

Case 3) When $(l-1)L + 2H < s < (l-2)L + 3H$, $\Pr(\xi \leq s) = (l+1)\beta^l - l\beta^{l+1} + \frac{(l+1)l}{2}\beta^{l-1}(1-\beta)^2$ and hence,

$$\begin{aligned}\frac{\partial}{\partial\beta}\Pr(\xi \leq s) &= (l+1)l\beta^{l-1}(1-\beta) + \frac{(l+1)l}{2}\left[(l-1)\beta^{l-2}(1-\beta)^2 - 2\beta^{l-1}(1-\beta)\right] \\ &= \frac{(l+1)l(l-1)}{2}\beta^{l-2}(1-\beta)^2 > 0.\end{aligned}$$

Case k) When $(l+1-k)L \leq s < (l-k)L + kH$ (where $k = 0, 1, ..., l$), the recursive form is

$$\frac{\partial}{\partial\beta}\Pr(\xi \leq s) = \frac{(l+1)!}{(l-k)!k!}\beta^{l-k}(1-\beta)^k > 0.$$

We conclude that $M(\beta, s)$ is strictly supermodular.

Let $s(\beta)$ be the firm's best response to β. The firm's objective is to minimize $M(\beta, s)$, and supermodularity implies that $s(\beta)$ is strictly increasing in β. Conversely, Nature's objective is to minimize $-M(\beta, s)$, indicating that $-\frac{\partial^2 M}{\partial\beta\partial s} < 0$ or submodularity. Let $\beta(s)$ be Nature's best response to s. The submodularity implies that $\beta(s)$ is strictly decreasing in s. We can visualize these two best response curves in a two-dimensional plane where the horizontal axis is β and the vertical axis is s. The firm's best response $s(\beta)$ moves from lower left to upper right, and Nature's best response $\beta(s)$ moves from upper left to lower right. The number of intercepts is either zero or one. Proposition 1 has established that the point (β^*, s_l^*) is an intercept of the two best response curves. Thus, the point (β^*, s_l^*) is the unique equilibrium (or intercept).

Proof of Proposition 2:

Using the steady-state distribution shown in equation (3.3), we find that the firm's time-average profit equals

$$\begin{aligned} M(\beta,s) &= \pi_1\left[pL - c(s-lL)\right] + \sum_{i=2}^{l} \pi_i\left(pL - cL\right) \\ &\quad + \left(\pi_{l+1}+\pi_{l+2}\right)\left\{-cL + \beta\left[pL - h\left(s-(l+1)L\right)\right] + (1-\beta)\,p\,(s-lL)\right\} \\ &= \frac{(1-\beta)\left[(lp+c)L - cs\right]}{l+1-l\beta} + \frac{\left[-cL + \beta\left[pL - h\left(s - lL - L\right)\right] + (1-\beta)\,p\,(s-lL)\right]}{l+1-l\beta} \\ &= \frac{\left[(p-c+(l+1)h)\beta L - \beta(p-c+h)s + (p-c)s\right]}{l+1-l\beta}. \end{aligned}$$

The FOC with respect to s yields

$$\frac{\partial M(\beta,s)}{\partial s} = \frac{-\beta(p-c+h)+(p+c)}{l+1-l\beta} = 0,$$

which confirms that $\beta^* = \frac{p-c}{p-c+h}$. We also find that under assumption (3.1),

$$\frac{\partial^2 M(\beta,s)}{\partial s \partial \beta} = \frac{-\beta(p-c+h)+(p+c)}{l+1-l\beta} = \frac{-(p-c)+lh}{(l+1+l\beta)^2} < 0,$$

which confirms submodularity.

The FOC with respect to β yields

$$\begin{aligned} \frac{\partial M(\beta,s)}{\partial \beta} &= -\frac{1}{(l+1-l\beta)^2}\left[(lp+c)L - cs\right] + \frac{1-\beta}{l+1-l\beta}\cdot\frac{\sigma(lp+c)}{2\beta\sqrt{\beta(1-\beta)}} \\ &\quad + \frac{l}{(l+1-l\beta)^2}\left[(l+1)(p+h)\beta L - \beta hs + (1-\beta)ps - lpL - cL\right] \\ &\quad + \frac{1}{l+1-l\beta}\left\{(l+1)(p+h)L - (p+h)s + \frac{\sigma\left[(l+1)(p+h)\beta - lp - c\right]}{2\beta\sqrt{\beta(1-\beta)}}\right\} \\ &= -\frac{p-c+(l+1)h}{(l+1-l\beta)^2}\left[s - (l+1)\mu + \frac{\sigma}{2}\frac{(l+1)-(l+2)\beta}{\sqrt{\beta(1-\beta)}}\right]. \end{aligned}$$

Given $\beta^* = \frac{p-c}{p-c+h}$, we find that

$$s^* = (l+1)\mu + \sigma\left(\frac{1}{2}\sqrt{\frac{p-c}{h}} - \frac{l+1}{2}\sqrt{\frac{h}{p-c}}\right),$$

which confirms equation (3.6). We also find that the first derivative with respect to l equals

$$\frac{\partial s^*}{\partial l} = \mu - \frac{\sigma}{2}\sqrt{\frac{h}{p-c}} \geq \mu - \frac{\sigma}{2\rho} = \frac{\mu}{2} > 0,$$

where we use the pre-condition (3.1). We conclude that the firm's robust optimal base-stock level s^* is increasing in l.

By substituting β^* and s^* into $M(\beta, s)$, we find that

$$Z^* = M(\beta^*, s^*) = \frac{1}{l+1-l\beta^*}\left[(p-c+(l+1)h)\beta^* L - \beta^*(p-c+h)s^* + (p-c)s^*\right]$$
$$= \frac{p-c+h}{p-c+(l+1)h}\begin{bmatrix}(p-c+(l+1)h)\frac{p-c}{p-c+h}(\mu - \sigma\sqrt{\frac{h}{p-c}}) \\ -\frac{p-c}{p-c+h}(p-c+h)s^* + (p-c)s^*\end{bmatrix} = (p-c)\mu - \sigma\sqrt{(p-c)h}.$$

The final step is to verify $(l+1)L^* \leq s^* \leq lL^* + H^*$. We find that

$$s^* - (l+1)L^* = (l+1)\,\mu + \sigma\left(\frac{1}{2}\sqrt{\frac{p-c}{h}} - \frac{l+1}{2}\sqrt{\frac{h}{p-c}}\right) - (l+1)\mu + (l+1)\sigma\sqrt{\frac{h}{p-c}}$$
$$= \frac{\sigma}{2}\left(\sqrt{\frac{p-c}{h}} + (l+1)\sqrt{\frac{h}{p-c}}\right) > 0,$$

and

$$lL^* + H^* - s^* = (l+1)\,\mu + \sigma\left(\sqrt{\frac{p-c}{h}} - l\sqrt{\frac{h}{p-c}}\right) - (l+1)\,\mu - \sigma\left(\frac{1}{2}\sqrt{\frac{p-c}{h}} - \frac{l+1}{2}\sqrt{\frac{h}{p-c}}\right)$$
$$= \frac{\sigma}{2}\sqrt{\frac{h}{p-c}}\left(\frac{p-c}{h} - l + 1\right) \geq 0,$$

where the last inequality is due to the assumption (3.1). Q.E.D.

Proof of Corollary 1

To compare the values of zero-sum games under backorder and lost sales, we first streamline the notation. Let $c_u = b = p - c$ be the understock cost and $c_o = h$ be the overstock cost. According to equation (2.7), the value of the zero-sum game under backorder equals

$$b\sigma(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}} = c_u\sigma(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}},$$

where $\beta^* = \left(\frac{c_u}{c_u+c_0}\right)^{\frac{1}{l+1}}$. The value of the zero-sum game under lost sales presented in Proposition 2 is the firm's profit, and we need to change it to the cost version. In particular, the mean of the firm's revenue is

$(p-c)E\left(\tilde{\theta}\right) = (p-c)\mu$; therefore, the value of the zero-sum game under lost sales in the cost minimization version equals

$$(p-c)\mu - \left((p-c)\mu - \sigma\sqrt{(p-c)h}\right) = \sqrt{c_u c_o}\sigma.$$

We find that when $l = 0$, $c_u\sigma(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}} = \sqrt{c_u c_o}\sigma$, indicating that backorder and lost sales are equally expensive.

The next task is to show that $(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}}$ is increasing in $l \geq 0$, which implies that when the lead time increases, backorders are strictly more expensive than lost sales. We transform the variables by letting $k = \frac{1}{l+1}$, $\alpha = \frac{c_u}{c_u+c_o}$, and $g(k) = \frac{1}{k}\sqrt{\frac{1-\alpha^k}{\alpha^k}}$, where $k \in (0,1]$ and $\alpha \in [\frac{1}{2}, 1)$. It holds that $g(k) = (l+1)\sqrt{\frac{1-\beta^*}{\beta^*}}$. The first derivative with respect to k equals

$$\frac{\partial g(k)}{\partial k} = -\frac{1}{k^2}\sqrt{\frac{1-\alpha^k}{\alpha^k}} - \frac{\ln\alpha}{2k\sqrt{\alpha^k(1-\alpha^k)}} = -\frac{\ln\alpha^k - 2\alpha^k + 2}{2k^2\sqrt{\alpha^k(1-\alpha^k)}}.$$

Let $w(x) = \ln x - 2x$, where $x = \alpha^k \in [\frac{1}{2}, 1)$ for $k \in (0,1]$. Because $\frac{\partial w(x)}{x} = \frac{1}{x} - 2 \leq 0$, $w(x)$ is decreasing in x. This means that $w(x) \geq w(1) = -2$, which further implies that $\ln\alpha^k - 2\alpha^k + 2 \geq 0$. Therefore, $\frac{\partial g(k)}{\partial k} \leq 0$. We conclude that $g(k)$ is decreasing in k and $(l+1)\sqrt{\frac{1-\beta^*}{\beta^*}}$ is increasing in l. Q.E.D.

Online Appendix: Nature's Full Model

The previous sections consider a relaxed version of the original zero-sum game. A crucial question is whether the relaxed solution indeed characterizes the true equilibrium. The key step is to investigate whether Nature changes its strategy when its action space expands from Ω_0 to Ω. If Nature does not alter its strategy, then the relaxed solution describes the true equilibrium.

Zero Lead Time

We start with the easiest case of zero lead time. To streamline the notations under backorders and lost sales, we define c_u as the understock cost and c_o as the overstock cost. Using the well-known identities $\min(\theta, q) = \theta - (\theta - q)^+ = q - (q - \theta)^+$, we find that the single-period ex post cost equals $Z(\theta, q) = c_u(\theta - q)^+ + c_o(q - \theta)^+$. Specifically, with lost sales, $c_u = p - c$ and $c_o = h$; and with backorders, $c_u = b$ and $c_o = h$. Sobel (1981) showed that a myopic policy is the equilibrium in a multiperiod stochastic game if the following three conditions hold: i) the reward depends on the current state and action; ii) the transition probability depends on the action but not on the current state; and iii) a static policy yields a repeatable Markov chain. With zero lead time, these three conditions hold; and thus, according to equation (2.3), we derive the myopic equilibrium by solving the following model:

$$Z = \min_{s \geq 0} \sup_{F \in \Omega} \left\{ \int_0^\infty c_u(\theta - s)^+ + c_o(s - \theta)^+ \, dF(\theta) \right\}, \tag{A-1}$$

which is indeed Scarf's model. It is well known that Nature prefers the two-point distributions defined in Ω_0. Therefore, with $L < s < H$ occurring in the equilibrium, the time-average cost equals

$$\begin{aligned} M(\beta, s) &= c_o\beta\left(s - \mu + \sigma\sqrt{\frac{1-\beta}{\beta}}\right) + c_u(1-\beta)\left(\mu + \sigma\sqrt{\frac{\beta}{1-\beta}} - s\right) \\ &= (c_o\beta + c_u\beta - c_u)(s - \mu) + (c_u + c_o)\sigma\sqrt{\beta(1-\beta)}, \end{aligned} \tag{A-2}$$

which yields the following results.

Corollary 2 *With $c_u \geq \rho^2 c_o$ and zero lead time, the two-player zero-sum game in equation (2.3) has a unique equilibrium satisfying*

$$\beta_0^* = \frac{c_u}{c_u + c_o} \text{ and } s_0^* = \mu + \frac{\sigma}{2}\left(\sqrt{\frac{c_u}{c_o}} - \sqrt{\frac{c_o}{c_u}}\right) \tag{A-3}$$

such that the value of the zero-sum game equals $Z^* = \sqrt{c_u c_o}\sigma$*. Nature's equilibrium strategy satisfies:*

$$F_0^* = \begin{cases} \Pr\left(\tilde{\theta} = \mu - \sigma\sqrt{\frac{c_o}{c_u}}\right) = \frac{c_u}{c_u + c_o}, \\ \Pr\left(\tilde{\theta} = \mu + \sigma\sqrt{\frac{c_u}{c_0}}\right) = \frac{c_o}{c_u + c_o}. \end{cases} \tag{A-4}$$

There are several noteworthy observations that can be made from Corollary 2. First, when $l = 0$, equations (2.6), (3.6), and (A-3) are identical, meaning that the firm's relaxed solution (under both backorder and lost sales) is optimal for the case with zero lead time. Second, when choosing the parameter β from the two-point distribution in Ω_0, Nature uses the firm's newsvendor ratio $\beta^* = \frac{c_u}{c_u + c_o}$. From the perspective of zero-sum games, Nature must anticipate the firm's best response to the chosen distribution. If the firm knows the demand distribution, she must apply her newsvendor ratio to determine her inventory level. Because the firm always applies her newsvendor ratio, Nature also uses the same ratio to determine its strategy. Hence, Nature's equilibrium strategy satisfies equation (A-4). Third, when $c_u \geq \rho^2 c_o$, the lower realized value of F_0^* in equation (A-4) is nonnegative, and the firm's equilibrium strategy happens to be the average of the two realized values of F_0^*.

Positive Lead Time and Backorder

We next study the backorder model with positive lead time. Clearly, the demand during the lead time $\tilde{\xi} = \sum_{i=1}^{l+1} \tilde{\theta}_i$ is a sufficient statistic to describe the system. Thus, after we replace $\left(\tilde{\theta}, F\right)$ with $\left(\tilde{\xi}, F_{l+1}\right)$, the three conditions in Sobel (1981) continue to hold, making the myopic equilibrium optimal. The remaining task is to solve the model in equation (2.3). A tempting shortcut is to apply Corollary 2 to construct a convolution F_{l+1}. However, we cannot identify an iid distribution F such that the corresponding convolution F_{l+1} can satisfy Corollary 2 unless $l = 0$. We must consider an alternative path to advance the analysis.

Suppose that Nature chooses a distribution F with $n \geq 3$ realized values rather than a two-point distribution. We denote $\lambda_i = Pr\left(\tilde{\theta} = \theta_i\right)$ for $i = 1, 2, ..., n$ as the probability when the realized demand is θ_i. To avoid degenerated solutions (otherwise, Nature chooses fewer points), we assume that $0 \leq \theta_1 < \theta_2 < ... < \theta_n$, $\lambda_1 > 0$, $\lambda_n > 0$ and $\lambda_i \geq 0$ for $i = 2, ..., n-1$. Before characterizing the equilibrium, we develop an important intermediate result as follows.

Lemma 1 *It holds that for* $i \in \{2, ..., n-1\}$,

$$\frac{\partial \theta_1}{\partial \theta_i} = -\frac{\lambda_i (\theta_n - \theta_i)}{\lambda_1 (\theta_n - \theta_1)} \text{ and } \frac{\partial \theta_n}{\partial \theta_i} = \frac{\lambda_i (\theta_1 - \theta_i)}{\lambda_n (\theta_n - \theta_1)}. \tag{A-5}$$

When Nature chooses a distribution with multiple points, Lemma 1 enables us to identify the relationship between the two end points (i.e., θ_1 and θ_n) and the middle points θ_i, where $i = 2, ..., n-1$. Specifically, we find that the first derivatives of both θ_1 and θ_n with respect to θ_i are linearly associated with λ_i. Based on this result, we can now derive the zero-sum game equilibrium when Nature chooses its strategy from Ω.

Proposition 3 *The firm's and Nature's strategies characterized in Proposition 2 are optimal for the original backorder model with a positive lead time.*

Proposition 3 indicates that when only the mean and variance of the demand are known, the firm's most unfavorable distribution is indeed a two-point distribution. Thus, the original backorder model is fully resolved. We next shift our attention to the case of lost sales.

Positive Lead Time and Lost Sales

We believe that the firm's relaxed solution in Proposition 2 is optimal for the lost-sales model with positive lead time. However, we are unable to produce a complete proof to show that $\lambda_i^* = 0$ for $i \in \{2, ..., n-1\}$. The curse of dimensionality prevents us from explicitly deriving the objective function. Nevertheless, to support our claim, we have managed to solve a special case where the lead time is $l = 1$ and Nature chooses a three-point distribution.

Corollary 3 *When the lead time* $l = 1$ *and Nature's strategy is a three-point distribution with parameters* (λ_i, θ_i)*, where* $\lambda_i = Pr\left(\tilde{\theta} = \theta_i\right)$ *for* $i \in \{1, 2, 3\}$ *and* $0 \le \theta_1 < \theta_2 < \theta_3$*, Nature's optimal strategy satisfies* $\lambda_2^* = 0$.

When $\lambda_2^* = 0$, Nature must prefer two-point rather than three-point distributions when $l = 1$. Corollary 3 can be extended to other cases where both n and l are small enough to be tractable (e.g., $(n, l) = (4, 2)$ is also tractable). For a generic pair of (n, l), the firm's time-average profit cannot be precisely derived; thus, we are unable to prove that Nature keeps its strategy unchanged. We hope that a new path to prove this conjecture can be discovered in future research.

Proof of Corollary 2

Using equation (A-2), we find the following two FOCs:

$$\frac{M(\beta,s)}{\partial s} = c_o\beta - c_u(1-\beta) = 0$$
$$\frac{M(\beta,s)}{\partial \beta} = (c_u+c_o)(s-\mu) + (c_u+c_o)\,\sigma\frac{(1-2\beta)}{2\sqrt{\beta(1-\beta)}} = 0,$$

yielding equation (A-3). We then substitute β_0^* and s_0^* into $M(\beta,s)$ to compute the value of the zero-sum game as follows:

$$Z^* = (c_u+c_o)\,\sigma\sqrt{\beta^*(1-\beta^*)} = (c_u+c_o)\,\sigma\sqrt{\frac{c_u}{c_u+c_o}\frac{c_o}{c_u+c_o}} = \sqrt{c_u c_o}\sigma.$$

This completes the proof of Corollary 2.   Q.E.D.

Proof of Lemma 1:

When Nature plays a distribution F with $n \geq 3$ realized values, its decision variables include (λ_i, θ_i), $i = 1, 2, ..., n$. The three moment conditions include:

$$\begin{cases} \lambda_1 + \sum_{k=2}^{n-1}\lambda_k + \lambda_n = 1, \\ \lambda_1\theta_1 + \sum_{k=2}^{n-1}\lambda_k\theta_k + \lambda_n\theta_n = \mu, \\ \lambda_1\theta_1^2 + \sum_{k=2}^{n-1}\lambda_k\theta_k^2 + \lambda_n\theta_n^2 = \mu^2 + \sigma^2. \end{cases}$$

We regard λ_1 and (λ_i, θ_i) as Nature's free decision variables (where $i = 2, ..., n-1$) and express the two end points (θ_1, θ_n) based on the other free variables. Specifically, the moment conditions yield that $\lambda_1 + \lambda_n = 1 - \sum_{k=2}^{n-1}\lambda_k$ and that

$$\theta_1 = \frac{\mu - \sum_{k=2}^{n-1}\lambda_k\theta_k}{\lambda_1+\lambda_n} - \sqrt{\frac{\lambda_n\sigma^2}{\lambda_1(\lambda_1+\lambda_n)} - \frac{\lambda_n Y}{\lambda_1(\lambda_1+\lambda_n)^2}},$$
$$\theta_n = \frac{\mu - \sum_{k=2}^{n-1}\lambda_k\theta_k}{\lambda_1+\lambda_n} + \sqrt{\frac{\lambda_1\sigma^2}{\lambda_n(\lambda_1+\lambda_n)} - \frac{\lambda_1 Y}{\lambda_n(\lambda_1+\lambda_n)^2}},$$

where

$$Y = \sum_{k=2}^{n-1}\lambda_k(\mu-\theta_k)^2 - \frac{1}{2}\sum_{k=2}^{n-1}\sum_{j=2}^{n-1}\lambda_k\lambda_j(\theta_k-\theta_j)^2.$$

Note that $\theta_k - \theta_j = 0$ whenever $k = j$. To have a better understanding of Y, consider a special case with $n = 4$, where

$$Y = \lambda_2 \left(\mu - \theta_2\right)^2 + \lambda_3 \left(\mu - \theta_3\right)^2 - \frac{1}{2}\lambda_2\lambda_3(\theta_2 - \theta_3)^2 - \frac{1}{2}\lambda_3\lambda_2(\theta_3 - \theta_2)^2.$$

For any $i = 2, \ldots, n-1$ (i.e., the middle $(n-2)$ points), it holds that

$$\frac{\partial\theta_1}{\partial\theta_i} = \lambda_i \left(\frac{-1}{\lambda_1 + \lambda_n} - \frac{\lambda_n(\mu - \theta_i) + \lambda_n \sum_{j=2}^{n-1} \lambda_j(\theta_i - \theta_j)}{\lambda_1 \left(\lambda_1 + \lambda_n\right)^2 \sqrt{\frac{\lambda_n}{\lambda_1(\lambda_1+\lambda_n)}\sigma^2 - \frac{\lambda_n}{\lambda_1(\lambda_1+\lambda_n)^2}Y}} \right).$$

Hence, $\frac{\partial\theta_1}{\partial\theta_i}$ takes the form of $\frac{\partial\theta_1}{\partial\theta_i} = \lambda_i K_{1,i}$. We continue to simplify $K_{1,i}$. According to the expressions of θ_1 and θ_n, we observe that

$$\begin{aligned}
\lambda_1\theta_1 &= \lambda_1 \frac{\mu - \sum_{k=2}^{n-1} \lambda_k\theta_k}{\lambda_1 + \lambda_n} - \sqrt{\frac{\lambda_1\lambda_n\sigma^2}{(\lambda_1 + \lambda_n)} - \frac{\lambda_1\lambda_n Y}{\left(\lambda_1 + \lambda_n\right)^2}}, \\
\lambda_n\theta_n &= \lambda_n \frac{\mu - \sum_{k=2}^{n-1} \lambda_k\theta_k}{\lambda_1 + \lambda_n} + \sqrt{\frac{\lambda_1\lambda_n\sigma^2}{(\lambda_1 + \lambda_n)} - \frac{\lambda_1\lambda_n Y}{\left(\lambda_1 + \lambda_n\right)^2}}.
\end{aligned}$$

Therefore, we find

$$\sqrt{\frac{\lambda_1\lambda_n\sigma^2}{(\lambda_1 + \lambda_n)} - \frac{\lambda_1\lambda_n Y}{\left(\lambda_1 + \lambda_n\right)^2}} = \lambda_1 \frac{\mu - \sum_{k=2}^{n-1} \lambda_k\theta_k}{\lambda_1 + \lambda_n} - \lambda_1\theta_1 = \lambda_n\theta_n - \lambda_n \frac{\mu - \sum_{k=2}^{n-1} \lambda_k\theta_k}{\lambda_1 + \lambda_n},$$

which implies that

$$\sqrt{\frac{\lambda_1\lambda_n\sigma^2}{(\lambda_1 + \lambda_n)} - \frac{\lambda_1\lambda_n Y}{\left(\lambda_1 + \lambda_n\right)^2}} = \frac{1}{2}\left(\lambda_n\theta_n - \lambda_1\theta_1 + \frac{(\lambda_1 - \lambda_n)(\lambda_1\theta_1 + \lambda_n\theta_n)}{\lambda_1 + \lambda_n} \right),$$

where we apply the mean constraint: $\mu - \sum_{k=2}^{n-1} \lambda_k\theta_k = \lambda_1\theta_1 + \lambda_n\theta_n$. We also find that

$$\mu - \theta_i + \sum_{j=2}^{n-1} \lambda_j(\theta_i - \theta_j) = \sum_{j=1}^{n} \left(\lambda_j\theta_j - \lambda_j\theta_i\right) + \sum_{j=2}^{n-1} (\lambda_j\theta_i - \lambda_j\theta_j) = \lambda_1 \left(\theta_1 - \theta_i\right) + \lambda_n \left(\theta_n - \theta_i\right).$$

After combining these equations, we find that

$$\begin{aligned}
K_{1,i} &= \frac{-1}{\lambda_1 + \lambda_n} - \frac{\lambda_n \left[\lambda_1 \left(\theta_1 - \theta_i\right) + \lambda_n \left(\theta_n - \theta_i\right)\right]}{\lambda_1 \left(\lambda_1 + \lambda_n\right)^2 \sqrt{\frac{\lambda_n}{\lambda_1(\lambda_1+\lambda_n)}\sigma^2 - \frac{\lambda_n}{\lambda_1(\lambda_1+\lambda_n)^2}Y}} \\
&= \frac{-1}{\lambda_1 + \lambda_n} \left(1 + \frac{\lambda_n \left[\lambda_1 \left(\theta_1 - \theta_i\right) + \lambda_n \left(\theta_n - \theta_i\right)\right]}{(\lambda_1 + \lambda_n) \sqrt{\frac{\lambda_1\lambda_n}{(\lambda_1+\lambda_n)}\sigma^2 - \frac{\lambda_1\lambda_n}{(\lambda_1+\lambda_n)^2}Y}} \right) \\
&= \frac{-1}{\lambda_1 + \lambda_n} \left(1 + \frac{2\lambda_n \left[\lambda_1 \left(\theta_1 - \theta_i\right) + \lambda_n \left(\theta_n - \theta_i\right)\right]}{(\lambda_1 + \lambda_n) \left(\lambda_n\theta_n - \lambda_1\theta_1 + \frac{(\lambda_1-\lambda_n)(\lambda_1\theta_1+\lambda_n\theta_n)}{\lambda_1+\lambda_n} \right)} \right) \\
&= \frac{-1}{\lambda_1 + \lambda_n} \left(1 + \frac{\lambda_n \left[\lambda_1 \left(\theta_1 - \theta_i\right) + \lambda_n \left(\theta_n - \theta_i\right)\right]}{\lambda_1\lambda_n(\theta_n - \theta_1)} \right) = \frac{-1}{\lambda_1 + \lambda_n} \frac{(\lambda_1 + \lambda_n)(\theta_n - \theta_i)}{\lambda_1(\theta_n - \theta_1)} = \frac{-\left(\theta_n - \theta_i\right)}{\lambda_1(\theta_n - \theta_1)},
\end{aligned}$$

which proves that $\frac{\partial\theta_1}{\partial\theta_i} = -\frac{\lambda_i(\theta_n-\theta_i)}{\lambda_1(\theta_n-\theta_1)}$.

Using the same method, we can also prove that $\frac{\partial\theta_n}{\partial\theta_i} = \lambda_i \frac{\theta_1-\theta_i}{\lambda_n(\theta_n-\theta_1)} = K_{n,i}$, where $K_{n,i} = \frac{\theta_1-\theta_i}{\lambda_n(\theta_n-\theta_1)}$. However, for any $k \notin \{1,i,n\}$, it holds that $\frac{\partial\theta_k}{\partial\theta_i} = 0$ because θ_k is freely chosen. Q.E.D.

Proof of Proposition 3:

The profile of Nature's n-point distribution is (θ_i,λ_i), in which θ_i is the realized value or point chosen by Nature and λ_i is the probability that Nature plays $\tilde{\theta} = \theta_i$ in any period, where $i \in \{1,2,...,n\}$. To illustrate the regular pattern, we first analyze a special case where $l = 1$ and then extend the analysis to the general case where $l \geq 1$. When $l = 1$, the firm's time-average cost equals

$$M = \sum_{t_1=1}^{n}\sum_{t_2=1}^{n} \lambda_{t_1}\lambda_{t_2}\left[b\left(\theta_{t_1}+\theta_{t_2}-s\right)^+ + h\left(s-\theta_{t_1}-\theta_{t_2}\right)^+\right], \tag{A-6}$$

where the vector $(\theta_{t_1},\theta_{t_2})$ represents the realized demand on a sample path during 2 periods. Here, $t_i \in \{1,2,...,n\}$ such that if $t_1 = 1$, the realized demand in period 1 is θ_1; and if $t_2 = 4$, the realized demand in period 2 is θ_4 (or the fourth point of Nature's distribution profile). There are n^2 pairs of (t_1,t_2) determining n^2 different sample paths. We let

$$T\left(t_1,t_2\right) = \lambda_{t_1}\lambda_{t_2}\left[b\left(\theta_{t_1}+\theta_{t_2}-s\right)^+ + h\left(s-\theta_{t_1}-\theta_{t_2}\right)^+\right] \tag{A-7}$$

be the weighted cost when the sample path is the vector $(\theta_{t_1},\theta_{t_2})$. The weight is the probability that the path occurs. When $\theta_{t_1}+\theta_{t_2} \geq s$, we find that $T(t_1,t_2) = \lambda_{t_1}\lambda_{t_2}b\left(\theta_{t_1}+\theta_{t_2}-s\right)$. Similarly, when $\theta_{t_1}+\theta_{t_2} < s$, we find that $T(t_1,t_2) = \lambda_{t_1}\lambda_{t_2}h\left(s-\theta_{t_1}-\theta_{t_2}\right)$.

We consider the case when the base-stock level satisfies $2\theta_1 < s < \theta_1+\theta_2$. Therefore, when $t_1 = t_2 = 1$, $T(1,1) = \lambda_1\lambda_1 h\left(s-\theta_1-\theta_1\right)$; however, for all the other $(t_1,t_2) \neq (1,1)$, $T(t_1,t_2) = \lambda_{t_1}\lambda_{t_2}b\left(\theta_{t_1}+\theta_{t_2}-s\right)$. We then find that

$$\begin{aligned} M &= \lambda_1\lambda_1 h\left(s-\theta_1-\theta_1\right) + \sum_{t_1=2}^{n}\sum_{t_2=2}^{n}\lambda_{t_1}\lambda_{t_2}b\left(\theta_{t_1}+\theta_{t_2}-s\right) \\ &= \lambda_1\lambda_1 h\left(s-\theta_1-\theta_1\right) - \lambda_1\lambda_1 b\left(\theta_1+\theta_1-s\right) + \sum_{t_1=1}^{n}\sum_{t_2=1}^{n}\lambda_{t_1}\lambda_{t_2}b\left(\theta_{t_1}+\theta_{t_2}-s\right) \\ &= \lambda_1\lambda_1\left(h+b\right)\left(s-\theta_1-\theta_1\right) + b\left(2\mu-s\right), \end{aligned}$$

where we apply the mean of Nature's demand profile $\sum_{k=1}^{n}\lambda_k\theta_k=\mu$. Hence, for any point $i\in\{2,...,n-1\}$ in the middle,

$$\frac{\partial M}{\partial\theta_i}=-2\lambda_1\lambda_1\left(b+h\right)\frac{\partial\theta_1}{\partial\theta_i}=-2\lambda_1\lambda_1\left(b+h\right)\left(-\frac{\lambda_i\left(\theta_n-\theta_i\right)}{\lambda_1(\theta_n-\theta_1)}\right)=\frac{2\lambda_1\lambda_i\left(b+h\right)\left(\theta_n-\theta_i\right)}{\left(\theta_n-\theta_1\right)},$$

where $\frac{\partial\theta_1}{\partial\theta_i}=-\frac{\lambda_i(\theta_n-\theta_i)}{\lambda_1(\theta_n-\theta_1)}$ follows Lemma 1. When $\lambda_1>0$ and $\theta_1<...<\theta_i<...\theta_n$, we conclude that if and only if $\lambda_i=0$, $\frac{\partial M}{\partial\theta_i}=0$. This directly implies that Nature chooses $\lambda_i^*=0$ for all the middle points $i\in\{2,...,n-1\}$. Therefore, for $l=1$, Nature's equilibrium strategy is a two-point distribution.

Next, we consider the general case where $l\geq 1$. Let $\left(\theta_{t_1},\theta_{t_2},...,\theta_{t_{l+1}}\right)$ be a vector with $(l+1)$ elements. Essentially, vector $\left(\theta_{t_1},\theta_{t_2},...,\theta_{t_{l+1}}\right)$ is an extension of vector $(\theta_{t_1},\theta_{t_2})$ as the lead time increases from 1 to l. There exist n^{l+1} sample paths. Using a similar method, we can verify that for $(l+1)\,\theta_1<s<l\theta_1+\theta_2$, the firm's time-average cost equals $M=(\lambda_1)^l\,(h+b)\,[s-(l+1)\,\theta_1]+b\,[(l+1)\,\mu-s]$ such that

$$\frac{\partial M}{\partial\theta_i}=\frac{(l+1)\,(\lambda_1)^l\,\lambda_i\,(b+h)\,(\theta_n-\theta_i)}{(\theta_n-\theta_1)}.$$

The FOC $\frac{\partial M}{\partial\theta_i}=0$ holds if and only if $\lambda_i^*=0$, for $i\in\{2,...,n-1\}$. We can now conclude that for any $l\geq 0$, Nature prefers two-point distributions from Ω_0. The relaxed solution in Proposition 1 is optimal for any lead time $l\geq 0$ in the original zero-sum game with backorder. Q.E.D.

Proof of Corollary 3:

When $l=1$, the order quantity placed in the period t arrives in the next period $t+1$. Similarly, we can describe the system by using $(q_t,q_{t-1};I_t)$, where the first variable q_t represents the order quantity that has just been placed in the current period t, the second variable q_{t-1} represents the delivery due in the current period, and the third variable represents the on-hand inventory before receiving the delivery q_{t-1}. It is more convenient to use $IN_t=q_{t-1}+I_t$ as the new state variable to replace (q_{t-1},I_t). The new variable IN_t represents the inventory available for sales. Under a base-stock level s, it must hold that $s=q_t+IN_t$. Because the base-stock level cannot be too high, it must hold that $IN_t<\theta_3$ (otherwise, stockouts never occurs in the steady state). We can verify that the Markov chain has 4 recurring states: $(\theta_1, s-\theta_1)$, $(\theta_2, s-\theta_2)$, $(s-\theta_2, \theta_2)$, and $(s-\theta_1, \theta_1)$, so that the Markov chain transition matrix is given as:

$$\begin{pmatrix} \lambda_1 & \lambda_2 & 0 & \lambda_3 \\ \lambda_1 & \lambda_2 & \lambda_3 & 0 \\ \lambda_1 & \lambda_2+\lambda_3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

We obtain the steady-state distributions as follows:

$$\pi_1 = \frac{1}{1+\lambda_3+\frac{(\lambda_2+\lambda_2\lambda_3)}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2}},\ \pi_2 = \frac{\lambda_2\pi_1}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2},\ \pi_3 = \lambda_3\pi_2,\ \text{and } \pi_4 = \lambda_3\pi_1. \tag{A-8}$$

We can then compute the time-average profit M in the steady state as follows:

$$\begin{aligned} M =& \ \pi_1\left[-c\theta_1 + \sum_{i=1}^{3}\lambda_i\left(p\min(\theta_i, s-\theta_1) - h(s-\theta_1-\theta_i)^+\right)\right] \\ &+\pi_2\left[-c\theta_2 + \sum_{i=1}^{3}\lambda_i\left(p\min(\theta_i, s-\theta_2) - h(s-\theta_2-\theta_i)^+\right)\right] \\ &+\pi_3\left[-c(s-\theta_2) + \sum_{i=1}^{3}\lambda_i\left(p\min(\theta_i, \theta_2) - h(\theta_2-\theta_i)^+\right)\right] \\ &+\pi_4\left[-c(s-\theta_1) + \sum_{i=1}^{3}\lambda_i\left(p\min(\theta_i, \theta_1) - h(\theta_1-\theta_i)^+\right)\right]. \end{aligned}$$

We explain the above equation as follows. In the steady state, the system is in state $(\theta_1, s-\theta_1)$ with probability π_1. In this state, the firm orders θ_1 units of inventory in the current period but has $s-\theta_1$ units of inventory available for sales. With probability λ_i, the realized demand in the current period is θ_i. The firm sells $\min(\theta_i, s-\theta_1)$ units of the product and keeps $(s-\theta_1-\theta_i)^+$ of leftover inventory.

We consider that $\theta_2 \le IN_t < \theta_3$ and simplify M as follows:

$$\begin{aligned} M =& \ \pi_1\begin{bmatrix} -c\theta_1 + \lambda_1(p\theta_1 - hs + 2h\theta_1) \\ +\lambda_2(p\theta_2 - hs + h\theta_1 + h\theta_2) + \lambda_3 p(s-\theta_1) \end{bmatrix} + \pi_2\begin{bmatrix} -c\theta_2 + \lambda_1(p\theta_1 - hs + h\theta_2 + h\theta_1) \\ +\lambda_2(p\theta_2 - hs + 2h\theta_2) + \lambda_3 p(s-\theta_2) \end{bmatrix} \\ &+ \pi_3\left[-c(s-\theta_2) + \lambda_1(p\theta_1 - h\theta_2 + h\theta_1) + \lambda_2 p\theta_2 + \lambda_3 p\theta_2\right] + \pi_4\left[-c(s-\theta_1) + \lambda_1 p\theta_1 + \lambda_2 p\theta_1 + \lambda_3 p\theta_1\right]. \end{aligned}$$

The FOC with respect to s equals:

$$\begin{aligned}\frac{\partial M}{\partial s} &= \pi_1\left[\lambda_3 p - h\left(\lambda_1+\lambda_2\right)\right] + \pi_2\left[\lambda_3 p - h\left(\lambda_1+\lambda_2\right)\right] + \pi_3\left(-c\right) + \pi_4\left(-c\right)\\ &= \left[\lambda_3 p - h\left(\lambda_1+\lambda_2\right)\right]\left(\pi_1+\pi_2\right) - c\left(\pi_3+\pi_4\right) - \left(\pi_1+\pi_2\right)\\ &= \left[\lambda_3 p - h\left(1-\lambda_3\right)\right]\left(\pi_1+\pi_2\right) - c\left(1-\pi_1-\pi_2\right) = \left(\lambda_3 p + \lambda_3 h - h + c\right)\left(\pi_1+\pi_2\right) - c = 0.\end{aligned}$$

On the other hand, equation (A-8) implies that

$$\pi_1 = \frac{1}{1+\lambda_3+\frac{(\lambda_2+\lambda_2\lambda_3)}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2}} \text{ and } \pi_2 = \frac{\lambda_2\pi_1}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2}.$$

We find that

$$\pi_1+\pi_2 = \frac{1}{1+\lambda_3+\frac{(\lambda_2+\lambda_2\lambda_3)}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2}}\left(1+\frac{\lambda_2}{1-\lambda_2-\lambda_2\lambda_3-\lambda_3^2}\right) = \frac{1}{1+\lambda_3}. \tag{A-9}$$

Thus, $\frac{\partial M}{\partial s}=0$ yields that

$$\frac{\lambda_3 p + \lambda_3 h - h + c}{1+\lambda_3} - c = 0,$$

indicating that $\lambda_3^* = \frac{h}{p-c+h}$.

The final step is to prove that $\lambda_2^* = 0$. We take the first derivative of M with respect to θ_2 and obtain the following result:

$$\begin{aligned}\frac{\partial M}{\partial \theta_2} &= \pi_1\begin{bmatrix} -c\frac{\partial\theta_1}{\partial\theta_2} + \lambda_1\left(p+2h\right)\frac{\partial\theta_1}{\partial\theta_2}\\ +\lambda_2\left(p+h\frac{\partial\theta_1}{\partial\theta_2}+h\right) - \lambda_3 p\frac{\partial\theta_1}{\partial\theta_2}\end{bmatrix} + \pi_2\begin{bmatrix} -c+\lambda_1\left(p\frac{\partial\theta_1}{\partial\theta_2}+h+h\frac{\partial\theta_1}{\partial\theta_2}\right)\\ +\lambda_2\left(p+2h\right)-\lambda_3 p\end{bmatrix}\\ &\quad +\pi_3\left[+c+\lambda_1\left(p\frac{\partial\theta_1}{\partial\theta_2}-h+h\frac{\partial\theta_1}{\partial\theta_2}\right)+\lambda_2 p+\lambda_3 p\right]\\ &\quad +\pi_4\left[+c\frac{\partial\theta_1}{\partial\theta_2}+\lambda_1 p\frac{\partial\theta_1}{\partial\theta_2}+\lambda_2 p\frac{\partial\theta_1}{\partial\theta_2}+\lambda_3 p\frac{\partial\theta_1}{\partial\theta_2}\right]\\ &= \pi_1\frac{\partial\theta_1}{\partial\theta_2}\left[-c+\lambda_1\left(p+2h\right)+\lambda_2 h-\lambda_3 p\right] + \left(\pi_2+\pi_3\right)\frac{\partial\theta_1}{\partial\theta_2}\left[\lambda_1\left(p+h\right)\right] + \pi_4\left(p+c\right)\frac{\partial\theta_1}{\partial\theta_2}\\ &\quad +\pi_1\lambda_2\left(p+h\right)+\pi_2\left(-c+\lambda_1 h+\lambda_2\left(p+2h\right)-\lambda_3 p\right)+\pi_3\left(c-\lambda_1 h+\lambda_2 p+\lambda_3 p\right).\end{aligned}$$

According to Lemma 1, it holds that $\frac{\partial\theta_1}{\partial\theta_2} = \lambda_2 K_{1,2}$. We obtain that

$$\begin{aligned}\frac{\partial M}{\partial \theta_2} &= \lambda_2 K_{1,2}\left\{\pi_1\left[-c+\lambda_1\left(p+2h\right)+\lambda_2 h-\lambda_3 p\right]+\left(\pi_2+\pi_3\right)\left[\lambda_1\left(p+h\right)\right]+\pi_4\left(p+c\right)\right\}\\ &\quad +\lambda_2\left[\pi_1\left(p+h\right)+\pi_2\left(p+2h\right)+\pi_3 p\right]+\left(\pi_2-\pi_3\right)\left(-c+\lambda_1 h-\lambda_3 p\right).\end{aligned}$$

In the above equation, the first two terms have a common factor λ_2. Because

$$\pi_2 - \pi_3 = \frac{\lambda_2 \pi_1}{1 - \lambda_2 - \lambda_2\lambda_3 - \lambda_3^2} - \lambda_3 \frac{\lambda_2 \pi_1}{1 - \lambda_2 - \lambda_2\lambda_3 - \lambda_3^2} = \frac{\lambda_2 (1 - \lambda_3) \pi_1}{1 - \lambda_2 - \lambda_2\lambda_3 - \lambda_3^2}.$$

The third term also has a common factor λ_2. By factoring out the common term λ_2, we find that

$$\begin{aligned} \frac{\partial M}{\partial \theta_2} &= \lambda_2 K_{1,2} \{\pi_1 [-c + \lambda_1 (p + 2h) + \lambda_2 h - \lambda_3 p] + (\pi_2 + \pi_3) [\lambda_1 (p + h)] + \pi_4 (p + c)\} \\ &\quad + \lambda_2 [\pi_1 (p + h) + \pi_2 (p + 2h) + \pi_3 p] + \lambda_2 \frac{(1 - \lambda_3) \pi_1 (-c + \lambda_1 h - \lambda_3 p)}{1 - \lambda_2 - \lambda_2\lambda_3 - \lambda_3^2}. \end{aligned}$$

By letting $\frac{\partial M}{\partial \theta_2} = 0$, we obtain an equilibrium where $\lambda_2^* = 0$. Accordingly, we find $\lambda_1^* = \frac{p-c}{p-c+h}$. Therefore, Nature's equilibrium strategy is a two-point distribution. Q.E.D.