Asymptotic Optimality of Order-Up-To Replenishment Policies for Serial Inventory Systems with Lost Sales

Marco Bijvank*  Woonghee Tim Huh†  Ganesh Janakiraman‡

Abstract

We study the optimal policy for a serial inventory system under periodic review when excess demand at the retailer (i.e., the most downstream stage) is lost. We focus on “high service level environments” (i.e., systems where the cost of a lost sale is high compared to inventory holding costs). These environments are typical of products whose margins are high relative to their holding costs. When excess demand is backordered, the optimal policy is a base-stock policy with base-stock levels calculated using the algorithm of Clark and Scarf [1960]. In this paper, we first propose to use this algorithm in lost-sales inventory systems by setting the backorder cost parameter in the algorithm to be equal to the lost-sales cost parameter. We show that the resulting base-stock policy is asymptotically optimal as the lost-sales penalty cost parameter grows. We also show that this result is robust in the following sense: There is a large family of choices for the backorder cost parameter used in the algorithm such that the asymptotic optimality continues to hold. Next, we propose a specific choice for the backorder cost parameter based on a power approximation formula. While the theoretical results guarantee asymptotical optimality as the service level approaches one, our computational investigation of problem instances with moderately high values for the service level (75%-99%) shows that the cost of the best base-stock policy is, on an average, 1.6% higher than the cost of the optimal policy, and the cost of the base-stock policy based on the power approximation is only 0.2% higher than that of the best base-stock policy.

Subject Classification: Inventory/Production: Lost Sales, Multi-echelon, Optimal Policy

*Haskayne School of Business, University of Calgary, Calgary, Canada. E-mail: marco.bijvank@haskayne.ucalgary.ca.
† Sauder School of Business, The University of British Columbia, Vancouver, BC Canada V6T 1Z4. E-mail: tim.huh@sauder.ubc.ca.
‡ School of Management, The University of Texas at Dallas, Richardson, TX 75080, USA. E-mail: ganesh@utdallas.edu.
1. Introduction

We study the problem of managing a serial inventory system under periodic review when excess demand at the retailer (i.e., the most downstream stage) is lost. Demands are stationary and stochastic. The cost model consists of a holding cost charged for each stage proportional to the amount of inventory at that stage and a penalty cost proportional to the amount of lost sales. We are interested in minimizing the long run average of the sum of the holding and penalty costs incurred per period. This problem is a building block for understanding how to manage multi-echelon, retail supply chains. In these supply chains, the assumption that excess downstream demand is lost is more meaningful than the assumption of backordering which pervades the multi-echelon inventory literature because the latter assumption lends analytical tractability to the problems.

If excess demand were completely backordered, we know that it is optimal to have all the locations follow echelon order-up-to policies [Clark and Scarf, 1960, Federgruen and Zipkin, 1984]. However, when excess demand is lost, even the optimal policy for a single location problem is not of such a simple form [Karlin and Scarf, 1958]. Therefore, it is clear that the optimal policy for the multi-echelon problem with lost sales will not be simple either. In this paper, we answer the following question: Under what conditions is the class of echelon order-up-to policies or echelon base-stock policies (used interchangeably by us) – policies that attempt to raise the total inventory in every echelon to a certain echelon-specific target level – “close to optimal”? We focus on “high service level environments” (i.e., systems where the cost of a lost sale is high compared to the inventory holding costs), and prove that echelon base-stock policies are asymptotically optimal as the penalty cost parameter grows.

To motivate the study of high service level environments, let us consider the following single-echelon example taken from Huh et al. [2009]: A retailer sells a product at a 25% mark-up, has a replenishment frequency of one week and an annual cost of capital of 15%; here, the ratio between the lost-sales cost to the holding cost is 100 which corresponds to an implied service level (i.e., the news-vendor ratio) of more than 99%. The parameters used in this example are quite typical of many retail products. Thus, for many retail product
categories, the service level is close to one. For a more detailed discussion on this issue, we refer the readers to Huh et al. [2009]. The other issue that deserves an explanation is the motivation for considering echelon base-stock policies. The most easily implementable class of policies in multi-echelon inventory systems is the class of installation or local base-stock policies – these policies raise the inventory position for every installation (or stage) of the supply chain to an installation-specific target level. It is easy to show that in our model, every stationary echelon base-stock policy can be implemented using a stationary installation base-stock policy [Axsäter and Rosling, 1993, Diks et al., 1996]. Since our model is a stationary model, the policies we consider are stationary echelon base-stock policies. Thus, they have the practical advantages of being implementable easily and locally. More specific details on our model follow in Section 3.

The policies we consider are the echelon base-stock policies suggested by the algorithm of Clark and Scarf [1960] or Federgruen and Zipkin [1984], originally shown to be optimal for backorder models. When we use this algorithm for lost-sales models, an input to this algorithm that we must specify is the backorder cost parameter. First, we show in Section 4 that when this parameter is chosen to be equal to the lost-sales cost parameter in our model, the resulting policy is asymptotically optimal. Our second result is the following generalization of the above result: We allow the backorder cost parameter input to the Clark-Scarf algorithm to be any increasing function of the lost-sales cost parameter whose slope is bounded below and above by strictly positive constants. We show that the asymptotic optimality result holds for this entire family of algorithms. This demonstrates the robustness of the asymptotic optimality phenomenon. These generalizations are presented in Section 5.

Our theoretical results open up two interesting questions which we also study: (a) How cost-effective are the best echelon base-stock policy at moderate service levels? (b) Given that there is a large family of asymptotically optimal echelon base-stock policies, how can we pick one which offers good performance across a wide range of problem parameters? To answer the latter question, we fit a power series with regression analysis on the best choice of the backorder cost parameter input as a function of the lost-sales cost parameter. This power approximation is developed in Section 6, and its performance is illustrated in Section 7 on
a numerical test bed with a wide range of parameter values when the lost-sales penalty cost parameter is moderately high, corresponding to implied service levels (defined as the newsvendor ratio at the most downstream stage) of 75% or more. The numerical experiments show that the policy computed using this approximation leads to a cost increase of less than 1% relative to the best base-stock policy. Furthermore, we perform a computational investigation of the best echelon base-stock policies, which provide excellent cost performance with an average gap of 1.6% relative to the costs of the optimal policies.

2. Related Literature

The existing literature in the five-decade-old field of inventory theory has seen only a few papers that study optimal policies for systems with lost sales and replenishment lead times and/or multiple inventory stages. This is despite the facts that the corresponding systems with backorders have been extensively studied and that the assumption of excess demand being lost is of as much practical importance as the backordering assumption. We will now briefly comment on the contributions of these few papers.

Karlin and Scarf [1958] consider a single location problem with a one-period lead time and show that there exists a critical amount of inventory below which it is optimal to order and above which it is optimal not to order. Furthermore, they show that the derivative of the optimal order quantity as a function of the inventory level is strictly between zero and −1 in the positive ordering region. Morton [1969] extends these results to the single location problem with arbitrary but deterministic integer lead times. He also derives lower and upper bounding functions on the optimal order policy. Zipkin [2008] has recently presented new, elegant proofs for Morton’s results. All of the papers mentioned above are for single-stage systems with lost sales. We refer to Bijvank and Vis [2011] for a recent literature overview on such inventory systems. For multi-echelon serial inventory systems, the only mathematical treatment of optimal policies, to our knowledge, is Huh and Janakiraman [2010]. This paper establishes the natural extensions of the single-stage results mentioned above.

The papers closest related to our work are Huh et al. [2009] and Bijvank et al. [2014]. They both study a single-stage inventory system with lost sales and arbitrary integer lead
times and show that the ratio between the optimal cost within the class of order-up-to policies to the optimal cost over all policies converges to one as the penalty cost approaches infinity. In practical terms, they show that a suitably chosen order-up-to policy is close to the optimal policy, in terms of cost, in high service level environments. Huh et al. [2009] show this asymptotic optimality result for a base-stock policy that is optimal in a backorder system with a particularly chosen backorder cost parameter. Bijvank et al. [2014] show that there exists a wide range of base-stock levels that can be used to accomplish this asymptotic optimality. We generalize the asymptotic result of Huh et al. [2009] to serial inventory systems for a specific value of the backorder cost parameter in Section 4. The robustness phenomenon of Bijvank et al. [2014] is extended to serial inventory systems in Section 5.

3. Notation

In this paper, we consider a multi-echelon serial inventory system with \( M \) stages, where we index stages by \( m = 1, \ldots, M \). External demand occurs at stage 1, stage \( m \) supplies stage \( m - 1 \) for all \( m \in \{2, \ldots, M\} \) and stage \( M \) is supplied by an external supplier with ample supply. The lead time between stage \( m + 1 \) and stage \( m \) is \( \tau_m \geq 1 \) periods. We index time periods in a forward manner, i.e., \( t = 1, 2, \ldots \). At the beginning of period \( t \), each stage \( m \) receives \( q_{m,t-\tau_m} \) units — this is the quantity stage \( m \) ordered \( \tau_m \) periods ago from stage \( m + 1 \) — and places a new replenishment order for \( q_{m,t} \) units from stage \( m + 1 \). At this instant, let \( x_{m,t} \) denote the amount of inventory on hand at stage \( m \). Next, external demand of \( D_t \) units is realized at stage 1, and demand is satisfied to the extent possible. We assume that demand that is not immediately satisfied is lost. Let \( e_{m,t} \) denote the amount of inventory in echelon \( m \) at this instant, i.e. the total amount of inventory held at all stages downstream of \( m \), including stage \( m \) and including all inventory in transit between these downstream stages. The echelon holding cost at stage \( m \) is \( h_m \) and the lost-sales penalty cost at stage 1 is \( p \). Demands are independently and identically distributed. We refer to this system as \( \mathcal{L}(h, \tau, p) \) or simply \( \mathcal{L} \) when there is no ambiguity. The performance measure of interest here is the long run average cost per period, i.e.,

\[
\limsup_{T \to \infty} \sum_{t=1}^{T} \left( E \left[ \sum_{m=1}^{M} h_m \cdot e_{m,t} + p \cdot (D_t - x_{1,t})^+ \right] \right) \frac{T}{T}.
\]
We prove our results by establishing relationships between $L$ and an identical multi-echelon system with backordering of excess demand, which we will denote by $B(h, \tau, b)$ or $B$. The parameter $b$ denotes the cost of backordering one unit of demand for one period. In these systems, we use $e_{m,t}$ to denote the *echelon m net-inventory* at the end of a period, i.e. the amount of physical inventory at stage $m$ and below minus the amount of backordered demand, if any, at stage 1. The performance measure we focus on for this system is

$$\limsup_{T \to \infty} \frac{\sum_{t=1}^{T} \left( E \left[ \sum_{m=1}^{M} h_m \cdot e_{m,t} + (b + h_1 + h_2 + \ldots + h_M) \cdot (e_{1,t})^- \right] \right)}{T},$$

where $(e_{1,t})^-$ is the amount of backordered demand at stage 1. This is a standard way of expressing the long run average cost in backorder systems. It can be verified algebraically that the expression within the expectation is the cost incurred in period $t$ by accounting for holding costs for every echelon charged proportionally to the amount of inventory on hand in that echelon and for backorder costs at stage 1 charged proportionally to the amount of backordered demand there. We know from Federgruen and Zipkin [1984] that an echelon base-stock policy is optimal for minimizing the long run average cost per period in this system. Let us use $S(b)$ to denote the vector of echelon order-up-to levels prescribed by the optimal policy for the backordering system, with the understanding that all parameters of the system other than $b$ are held constant. Note that $S(b)$ can be computed using the algorithm described in Federgruen and Zipkin [1984], which is based on the algorithm of Clark and Scarf [1960] for the finite horizon problem.

Our main theoretical result is that, as $p$ becomes arbitrarily large, an asymptotically optimal policy for the lost-sales system $L(h, \tau, p)$ is given by an echelon base-stock policy whose echelon order-up-to levels are given by the vector $S(b_p)$, where

$$b_p = p + \sum_{m=1}^{M} h_m \cdot \sum_{j=m}^{M} \tau_j. $$

The asymptotic optimality result is formally stated and proven in the next section. Moreover, we show in Section 5 that this asymptotic optimality is robust in the sense that it holds for a wide choice of echelon order-up-to vectors.
4. Asymptotic Optimality of Echelon Order-Up-To Policies

Let $D$ denote the random variable representing the single period demand. Let $C^{L,*}(h, \tau, p)$ and $C^{B,*}(h, \tau, b)$ denote the long run average costs for the optimal policy in $L$ and $B$, respectively. Let $C^{L,S}(h, \tau, p)$ and $C^{B,S}(h, \tau, b)$ denote the long run average costs under the echelon base-stock policy with order-up-to levels $S$ in $L$ and $B$, respectively. From the optimality of base-stock policies in $B$ and by definition of $S(b)$, we have the identity

$$C^{B,*}(h, \tau, b) = C^{B,S(b)}(h, \tau, b).$$

(1)

Let $D[t_1, t_2] = \sum_{t=t_1}^{t_2} D_t$ represent the cumulative demand incurred in the interval $[t_1, t_2]$.

**Theorem 1.** Assume that $D$ is an unbounded random variable, i.e. $\sup\{x : P(D \leq x) < 1\} = \infty$, and that $E[D] < \infty$. Also, assume that for any positive integer $k$,

$$\lim_{d \to \infty} E \left[ D[1,k] - d \mid D[1,k] > d \right] / d = 0.$$

Then,

$$\lim_{p \to \infty} \left( \frac{C^{L,S(b_p)}(h, \tau, p)}{C^{L,*}(h, \tau, p)} \right) = 1.$$

Note that Theorem 1 requires a technical assumption that the demand distribution must satisfy $E \left[ D[1,k] - d \mid D[1,k] > d \right] / d \to 0$ as $d \to \infty$. This assumption appeared in Huh et al. [2009]. They explain that many commonly used distributions satisfy this assumption including geometric distributions, Poisson distributions, negative binomial distributions with parameters $r > 0$ and $0 < \rho < 1$, exponential distributions, and Gaussian distributions. Theorem 1 also holds when $D$ is bounded, which we will discuss in Section 5.2.

### 4.1 Proof of Theorem 1

We prove Theorem 1 for a two-echelon system through a sequence of preliminary lemmas. It will be obvious that all the proofs in this section extend directly to multi-echelon serial systems, in general. Thus, in the remainder of this section, we assume $M = 2$. The proof of Theorem 1 consists of three parts: (1) Developing a lower bound on the denominator in the statement of the theorem, $C^{L,*}(h, \tau, p)$, (2) developing an upper bound on the numerator,
$C^L_S(b_p)(h, \tau, p)$, and then (3) showing that the ratio between the upper bound and the lower bound approaches 1. While this sequence of steps is identical to that of Huh et al. [2009] who study single-stage systems, the proofs of these steps are more involved in the multi-echelon case studied here. We comment on this issue at the end of this section.

The following lemma shows that $C^L_S(h, \tau, p)$ is bounded below by the optimal cost of a backordering system with a suitably defined backorder cost. The proofs for the results in this section are presented in Appendix A.

**Lemma 2.** The following inequality holds for all $(h, \tau, p)$:

$$C^L_S(h, \tau, p + \sum_{m=2}^{M} h_m \cdot \sum_{j=1}^{m-1} \tau_j) \geq C^{B,*}(h, \tau, p/(\tau_1 + \tau_2 + 1)).$$

This implies that $C^L_S(h, \tau, p) \geq C^{B,*}(h, \tau, (p - \sum_{m=2}^{M} h_m \cdot \sum_{j=1}^{m-1} \tau_j)/(\tau_1 + \tau_2 + 1))$ holds for all $(h, \tau, p)$.

Now, we proceed to establish an upper bound for $C^L_S(b_p)(h, \tau, p)$, the numerator in the statement of Theorem 1. We will obtain this result as a corollary of the following lemma.

**Lemma 3.** For any S, $C^L_S(h, \tau, p) \leq C^{B,S}(h, \tau, b_p)$.

As a corollary of Lemma 3, we establish the following upper bound on $C^L_S(b_p)(h, \tau, p)$.

It follows from the fact that $C^{B,S}(h, \tau, b_p)$ is equal to $C^{B,*}(h, \tau, b_p)$ by (1).

**Corollary 4.** $C^L_S(b_p)(h, \tau, p) \leq C^{B,*}(h, \tau, b_p)$.

So far, in this section, we have established a lower bound on the denominator in the statement of the theorem (Lemma 2) and an upper bound on the numerator (Corollary 4). For the remainder of this section, we study their ratio. To do this, we first present a result on the backordering system.

**Lemma 5.** Under the conditions of Theorem 1, for any $\beta > 0$,

$$\lim_{b \to \infty} \left( \frac{C^{B,*}(h, \tau, \beta \cdot b)}{C^{B,*}(h, \tau, b)} \right) = 1.$$
We are now ready to achieve the main purpose of this section which is to prove Theorem 1. Recall that $b_p = p + h_1 \cdot (\tau_1 + \tau_2) + h_2 \cdot \tau_2$. It is easy to observe that Lemma 2 and Corollary 4 imply the following:

$$\frac{C^{CL, S(b_p)}(h, \tau, p)}{C^{CL, S}(h, \tau, p)} \leq \frac{C^{B, *}(h, \tau, b_p)}{C^{B, *} (h, \tau, p/(\tau_1 + \tau_2 + 1))} \leq \frac{C^{B, *}(h, \tau, 2p)}{C^{B, *} (h, \tau, p/(\tau_1 + \tau_2 + 1))},$$

for sufficiently large values of $p$. Lemma 5 implies that the rightmost expression above approaches 1 as $p \to \infty$. This completes the proof of Theorem 1.

We conclude this section with some comments comparing the proofs in this section with the proofs of corresponding results in Huh et al. [2009]. The proofs of Lemmas 2 and 3 (deriving a lower bound on the optimal cost and an upper bound on the cost of an echelon order-up-to policy) involve comparing the dynamics in a lost sales inventory system with those in a backorder system. The task of making this comparison across every echelon as opposed to a single echelon comparison makes these proofs more difficult than the proofs in Huh et al. The bounds of Lemmas 2 and 3 are expressed as the optimal costs of two appropriately defined multi-echelon, backorder systems. Lemma 5 presents a bound on the ratio of these two optimal costs. In the single echelon case studied by Huh et al., the optimal cost in a backorder system is identical to the optimal cost in a newsvendor problem – this fact expedited their analysis of the above ratio. In the multi-echelon case, however, the optimal cost for a backorder system does not have a closed-form or a newsvendor-type expression. Consequently, we use Shang and Song [2003]'s newsvendor-type bounds on the optimal cost of a multi-echelon backorder system to prove Lemma 5.

## 5. Generalization of Theorem 1

Theorem 1 shows the asymptotic optimality of an echelon order-up-to policy as $p \to \infty$. The particular choice of the vector of echelon base-stock levels is $S(b_p)$ where

$$b_p = p + \sum_{m=1}^{M} h_m \cdot \sum_{j=m}^{M} \tau_j.$$

In Section 5.1, we establish the robustness of this asymptotic optimality by demonstrating that there is an entire family of echelon order-up-to vectors which also achieve this asymptotic optimality. This robustness result was not previously noted by Huh et al. [2009] and is new.
even for the single stage system, and it is one of the contributions of this paper. Furthermore, we prove in Section 5.2 that Theorem 1 and the above-mentioned robustness result also hold when the demand is bounded.

5.1 Robustness of the Asymptotic Optimality

In this section, we show that the optimality result of Theorem 1 holds for any selection of the order-up-to vector in the following form: \( S(\bar{b}_p) \), where \( \bar{b}_p = \theta(p) \) for any function \( \theta \) satisfying Assumption 1 below.

**Assumption 1.** There exists a triplet \((p_0, \theta, \bar{\theta})\), where \( p_0 < \infty \) and \( 0 < \theta < \bar{\theta} < \infty \), such that

\[
\theta \cdot b_p \leq \theta(p) \leq \bar{\theta} \cdot b_p \quad \text{for all } p \geq p_0.
\]

We first generalize Lemma 5.

**Lemma 6.** Consider any function \( \beta(b) \) such that, for some \( 0 < \beta < \infty \), \( \beta > 0 \) and \( b_0 < \infty \), we have \( \beta(b) \geq \beta \cdot b \) for all \( b \) and \( \beta(b + \epsilon) - \beta(b) \leq \bar{\beta} \cdot \epsilon \) for all \( b \geq b_0 \) and all \( \epsilon > 0 \). Then, under the conditions of Theorem 1,

\[
\lim_{b \to \infty} \left( \frac{C^B^*(h, \tau, \beta(b))}{C^B^*(h, \tau, b)} \right) = 1.
\]

**Proof.** Let \( b \geq \max\{b_0, \beta(b_0)/\bar{\beta}\} \). Then, the conditions imposed on \( \beta(\cdot) \) in the statement of the lemma imply that

\[
\beta(b) \leq \beta(b_0) + \bar{\beta} \cdot (b - b_0) \leq \beta(b_0) + \bar{\beta} \cdot b \leq 2 \cdot \bar{\beta} \cdot b.
\]

Thus, \( C^B^*(h, \tau, \beta(b)) \) is bounded above by \( C^B^*(h, \tau, 2 \cdot \bar{\beta} \cdot b) \). Moreover, since \( \beta(b) \geq \beta \cdot b \), we know that \( C^B^*(h, \tau, \beta(b)) \) is bounded below by \( C^B^*(h, \tau, \beta \cdot b) \). Therefore,

\[
\left( \frac{C^B^*(h, \tau, \beta \cdot b)}{C^B^*(h, \tau, b)} \right) \leq \left( \frac{C^B^*(h, \tau, \beta(b))}{C^B^*(h, \tau, b)} \right) \leq \left( \frac{C^B^*(h, \tau, 2 \cdot \bar{\beta} \cdot b)}{C^B^*(h, \tau, b)} \right).
\]

Since both the leftmost expression and the rightmost expression in (2) converge to 1 as \( b \to \infty \) by Lemma 5, we conclude that the middle expression also converges to 1.
Theorem 7. Under the conditions of Theorem 1,

$$\lim_{p \to \infty} \left( \frac{C^{L,S(\theta(p))}(h, \tau, p)}{C^{L,*}(h, \tau, p)} \right) = 1,$$

for any function \( \theta(\cdot) \) satisfying Assumption 1.

Proof. As in the proof of Theorem 1, we focus our proof on the two-echelon case, i.e., \( M = 2 \). We know from Lemma 2 that

$$C^{L,*}(h, \tau, p) \geq C^{B,*}(h, \tau, (p - h_2 \cdot \tau_1)/(\tau_1 + \tau_2 + 1)).$$

For any \( p \geq p_0 \), we have \( \theta(p) \geq \theta \cdot b_p \) by Assumption 1. Using this fact along with Lemma 3, we obtain

$$C^{L,S(\theta(p))}(h, \tau, p) \leq C^{B,S(\theta(p))}(h, \tau, b_p) \leq C^{B,S(\theta(p))}(h, \tau, \theta(p)/\theta).$$

Thus, from the above two sets of inequalities, it suffices to show that

$$\lim_{p \to \infty} \frac{C^{B,S(\theta(p))}(h, \tau, \theta(p)/\theta)}{C^{B,*}(h, \tau, (p - h_2 \cdot \tau_1)/(\tau_1 + \tau_2 + 1))} = 1,$$

which implies the required result. We can write the ratio within the limit of (3) as the following product

$$\left( \frac{C^{B,S(\theta(p))}(h, \tau, \theta(p)/\theta)}{C^{B,*}(h, \tau, \theta(p)/\theta)} \right) \cdot \left( \frac{C^{B,*}(h, \tau, \theta(p)/\theta)}{C^{B,*}(h, \tau, (p - h_2 \cdot \tau_1)/(\tau_1 + \tau_2 + 1))} \right) \quad (4)$$

It is sufficient to show that both the ratios in (4) converge to 1.

We first consider the second ratio in (4). Define

$$\beta(u) = \theta(h_2 \cdot \tau_1 + (\tau_1 + \tau_2 + 1) \cdot u)/\theta \quad \text{and} \quad \hat{b}(p) = (p - h_2 \cdot \tau_1)/(\tau_1 + \tau_2 + 1).$$

From the above definitions, it follows that \( \theta(p)/\theta = \beta(\hat{b}(p)) \), and the second ratio in (4) equals \( C^{B,*}(h, \tau, \beta(\hat{b}(p))) / C^{B,*}(h, \tau, \hat{b}(p)) \). As \( p \) approaches \( \infty \), so does \( \hat{b}(p) \). Also, it is trivial to verify that the function \( \beta(\cdot) \) defined above satisfies the conditions of Lemma 6. Thus, the limit of this ratio equals 1 by that lemma.

We now consider the first ratio of (4), which equals \( C^{B,S(\theta(p))}(h, \tau, \hat{b}(p)) / C^{B,*}(h, \tau, \hat{b}(p)) \) by defining \( \hat{b}(p) = \theta(p)/\theta \). We want to show that this ratio converges to 1. Next, notice
from Assumption 1 that there is no loss of generality in assuming \( \theta \leq 1 \). Thus, \( \hat{b}(p) \geq \theta(p) \).

This implies that

\[
\frac{C^{B,S(\theta(p))}(h, \tau, \hat{b}(p))}{C^{B,*}(h, \tau, \hat{b}(p))} \leq \frac{C^{B,S(\theta(p))}(h, \tau, \hat{b}(p))}{C^{B,*}(h, \tau, \theta(p))}.
\]

Since the ratio on the left is greater than or equal to 1, it is sufficient to show that the ratio on the right converges to 1.

As in Section 4.1, let \( BO_b \) denote the expected amount on backorder under an echelon order-up-to policy with echelon base-stock levels \( S_1(b) \) and \( S_2(b) \). Let us now compare the costs incurred by the two backorder systems, \( B(h, \tau, \hat{b}(p)) \) and \( B(h, \tau, \theta(p)) \) (which is the same as \( B(h, \tau, \theta \cdot \hat{b}(p)) \) from the definition of \( \hat{b}(p) \)) when they are both managed by the echelon order-up-to \( S(\theta(p)) \) policy (this is the optimal policy for the latter system). The dynamics of these two systems are the same. Thus the cost difference is due to backordering costs only, and is given by

\[
C^{B,S(\theta(p))}(h, \tau, \hat{b}(p)) - C^{B,*}(h, \tau, \theta(p))
= C^{B,S(\theta(p))}(h, \tau, \hat{b}(p)) - C^{B,S(\theta(p))}(h, \tau, \theta(p))
= C^{B,S(\theta(p))}(h, \tau, \hat{b}(p)) - C^{B,S(\theta(p))}(h, \tau, \theta \cdot \hat{b}(p))
= (1 - \theta) \cdot \hat{b}(p) \cdot BO_{\hat{b}(p)}.
\]

Therefore,

\[
\frac{C^{B,S(\theta(p))}(h, \tau, \hat{b}(p))}{C^{B,*}(h, \tau, \theta(p))} = 1 + \frac{(1 - \theta) \cdot \hat{b}(p) \cdot BO_{\hat{b}(p)}}{C^{B,*}(h, \tau, \theta \cdot \hat{b}(p))}.
\]

It is sufficient to show the claim that the quantity on the right converges to 1. This quantity can be rewritten as

\[
1 + \left( \frac{1 - \theta}{\theta} \right) \cdot \left[ \frac{\theta \cdot \hat{b}(p) \cdot BO_{\hat{b}(p)}}{C^{B,*}(h, \tau, \theta \cdot \hat{b}(p))} \right].
\]

As \( p \to \infty \), \( \theta \cdot \hat{b}(p) \) also approaches \( \infty \). By using (20), we know that the expression in the square bracket approaches 0. This immediately implies the claim above, which completes the proof of the theorem. \( \square \)
5.2 Bounded Demand Distribution

In the statements of Theorem 1 and Theorem 7, we have required that the demand distribution $D$ is unbounded and that it satisfies a technical condition involving its tail distribution. These assumptions are used in the proof of Lemma 5, in particular, to prove the claim in (20). We show that the results of these theorems hold even when $D$ is bounded. To accomplish this, it suffices to show that the claim in (20) holds when $D$ is bounded.

For convenience, we present (20) below:

$$\lim_{b \to \infty} \frac{b \cdot BO_b}{C^{B,*}(h, \tau, b)} = 0.$$  

Since the denominator of the ratio above is bounded below by a strictly positive constant for sufficiently large $b$, it suffices to show that the numerator converges to zero as $b \to \infty$. To see this, define two events $E_1 = [D[\tau_2 + 1, \tau_1 + \tau_2 + 1] > S_1(b)]$ and $E_2 = [D[1, \tau_1 + \tau_2 + 1] > S_2(b)]$. Then, from (21),

$$b \cdot BO_b \leq b \cdot E \left[ (D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+ \right] + b \cdot E \left[ (D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+ \right]$$

$$= b \cdot P[E_1] \cdot E \left[ (D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+ \mid E_1 \right]$$

$$+ b \cdot P[E_2] \cdot E \left[ (D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+ \mid E_2 \right].$$

From (18) and (19), we obtain that

$$b \cdot P[E_1] \leq b \cdot h_1/(b + h_1 + h_2) \leq h_1 \quad \text{and} \quad$$

$$b \cdot P[E_2] \leq b \cdot (h_1 + h_2)/(b + h_1 + h_2) \leq (h_1 + h_2).$$

Moreover, the above conditional expectation expressions converge to 0 as $b \to \infty$.\(^1\) This proves the claim in (20).

\(^1\)To prove this convergence, let $\gamma = \sup\{x : P(D < x) < 1\}$, i.e., $\gamma$ is the right-end of the support of $D$. Then, as $b \to \infty$, we know that $S_1(b) \to \gamma \cdot (\tau_1 + 1)$ and $S_2(b) \to \gamma \cdot (\tau_1 + \tau_2 + 1)$. By definition of $\gamma$, we also have $D[\tau_2 + 1, \tau_1 + \tau_2 + 1] \leq \gamma \cdot (\tau_1 + 1)$ with probability 1 and $D[1, \tau_1 + \tau_2 + 1] \leq \gamma \cdot (\tau_1 + \tau_2 + 1)$ with probability 1. This implies the convergence of the conditional expectations to zero.
6. Power Approximation

In Section 5, we proved that there is a large family of functions \( \theta(p) \) for which the corresponding echelon base-stock levels \( S(\theta(p)) \) are asymptotically optimal for large values of the penalty cost \( p \). However, from a practical perspective, it is important to identify an easily computable function \( \theta(p) \) from this family, which provides excellent cost performance across a wide range of problem instances including those with moderately high service levels (or equivalently, moderately high values of \( p \)). This is the subject of this section.

Let us first discuss some desirable properties that seem reasonable to impose on the function \( \theta(p) \).

1. For the special case of a single echelon and a zero period lead time, the inventory problem we study is nothing but the multi-period newsvendor problem. In this case, we know that the optimal policy is to order up to \( S(p) \) in every period. So, we require \( \theta(p) \) to be the identity function (i.e., \( \theta(p) = p \)) in this case.

2. It is desirable that our choice of \( \theta(p) \) satisfies Assumption 1 so that the order-up-to \( S(\theta(p)) \) policy is also endowed with the asymptotic optimality property of Section 5.

3. Let

\[
\tilde{p} = \left( p - \sum_{m=2}^{M} h_m \cdot \sum_{j=1}^{m-1} \tau_j \right) / \left( \sum_{i=1}^{M} \tau_i + 1 \right),
\]

and let \( S^*(p) \) denote the best vector of order-up-to levels in \( \mathcal{L}(h, \tau, p) \). From Lemma 2 and Lemma 3, we can bound the optimal cost \( C_{\mathcal{L},S^*}(p)(h, \tau, p) \) below by

\[
C_{\mathcal{L},S^*}(p)(h, \tau, p) \geq C_{\mathcal{L},S^*}(h, \tau, p) \geq C_{\mathcal{B},S}(\tilde{p})(h, \tau, \tilde{p}),
\]

and above by

\[
C_{\mathcal{L},S^*}(p)(h, \tau, p) \leq C_{\mathcal{L},S}(b_p)(h, \tau, p) \leq C_{\mathcal{B},S}(b_p)(h, \tau, b_p).
\]

This motivates the conjecture that the optimal choice of \( \theta(p) \) for any given problem instance lies between \( \tilde{p} \) and \( b_p \). We observed from our computational experiments that
this statement holds in the majority of the problem instances (the optimal choice of 
\( \theta(p) \) sits outside this range in less than 7% of the instances in our experiment). As a 
result of this observation, we restrict our attention to the case where \( \theta(p) \) is a convex 
combination of \( \tilde{\theta} \) and \( \theta_p \).

In particular, we let the weights used in the convex combination depend on the implied 
service level \( SL = p/(p + \sum_i h_i) \) only:

\[
\theta(p) = \left( \alpha(SL) \tilde{\theta} + (1 - \alpha(SL)) \theta_p \right)^+, \quad 0 \leq \alpha(SL) \leq 1.
\] (5)

(Note that the positive part operator is to ensure that \( \theta(p) \) is nonnegative since \( \tilde{\theta} \) can be 
negative when \( p \) is small.) The reason for choosing the above form of \( \theta(p) \) is twofold. First, it 
satisfies all three desirable properties discussed above. Since \( \alpha(SL) \) is defined on the interval 
\([0, 1]\), it is easy to show that \( \theta(p) \) satisfies Assumption 1. Second, it reduces the search for 
\( \theta(p) \), which can be a function of \( h, p, \tau \) and the demand distribution (while the notation \( \theta(p) \) 
suggests the dependence on \( p \) alone, the dependence on the other parameters is implicitly 
assumed), to a search for \( \alpha \), which is a function of the implied service level alone. It still 
remains to specify a form for the function \( \alpha(SL) \).

Let \( \underline{b}(h, \tau, p) \) be the smallest value of \( b \) minimizing \( C_{L,S}(b)(h, \tau, p) \), and similarly let 
\( \overline{b}(h, \tau, p) \) be the largest value of \( b \) minimizing \( C_{L,S}(b)(h, \tau, p) \). Because of the monotonicity 
of \( S(b) \) in \( b \) [Shang and Song, 2003], any value \( b \in [\underline{b}(h, \tau, p), \overline{b}(h, \tau, p)] \) minimizes 
\( C_{L,S}(b)(h, \tau, p) \). Motivated by (5), we define

\[
\underline{\alpha}(h, \tau, p) = \frac{\overline{b}(h, \tau, p) - b_p}{\tilde{\theta} - b_p},
\] (6)

\[
\overline{\alpha}(h, \tau, p) = \frac{\underline{b}(h, \tau, p) - b_p}{\tilde{\theta} - b_p}.
\] (7)

Note that it would be ideal if our choice of the \( \alpha(SL) \) function in \( SL = p/(p + \sum_i h_i) \) 
belongs to the interval \([\underline{\alpha}(h, \tau, p), \overline{\alpha}(h, \tau, p)]\), since it would imply that our choice of \( \alpha \) 
would produce the best choice for the backward penalty parameter for the algorithm. Next, 
we use a numerical analysis to fit a function to \( \alpha(SL) \) with a regression model using the 
interval \([\underline{\alpha}(h, \tau, p), \overline{\alpha}(h, \tau, p)]\) of each instance from the data.
The data used for this numerical analysis is similar to the numerical setting used by Shang and Song [2003]. Table 1 lists the parameter settings. Demand is generated from a Poisson distribution with three mean values per period, 4, 8, and 16. We consider a setting with 2 and 4 echelon levels, where the total system lead time equals 4 periods. That is, \( \tau_i = 1 \) when \( M = 4 \), and \( \tau_i = 2 \) for \( M = 2 \). The cost function is linear in the parameters \( h_i \) and \( p \), where the echelon unit holding costs are the same across all echelons and chosen from \( \{0.25, 0.5, 1, 2\} \), and the unit penalty cost is one of \( \{9, 19, 39, 59, 79, 99\} \). All combinations of these parameter settings result in 144 test instances. The implied service level for the settings given in Table 1 ranges from 52.94% to 99.50%.

<table>
<thead>
<tr>
<th>parameter</th>
<th>values</th>
</tr>
</thead>
<tbody>
<tr>
<td>demand distribution</td>
<td>Poisson</td>
</tr>
<tr>
<td>mean demand</td>
<td>4, 8, 16</td>
</tr>
<tr>
<td>echelon levels ((M))</td>
<td>2, 4</td>
</tr>
<tr>
<td>replenishment lead time ((\tau_i))</td>
<td>1 when ( M = 2 )</td>
</tr>
<tr>
<td></td>
<td>2 when ( M = 4 )</td>
</tr>
<tr>
<td>unit holding cost ((h_i))</td>
<td>0.25, 0.5, 1, 2</td>
</tr>
<tr>
<td>unit penalty cost ((p))</td>
<td>9, 19, 39, 59, 79, 99</td>
</tr>
</tbody>
</table>

Table 1: System parameters for the problem instances to determine \( \alpha(SL) \).

Our experiments suggest an exponential or power form for \( \alpha(SL) \). More specifically, we focus on the following form,

\[
\alpha(SL) = 1 - c_0(SL)^{c_1}, \tag{8}
\]

where \( c_0 \) and \( c_1 \) are constants to be estimated with non-linear regression. Note that \( 0 \leq \alpha(SL) \leq 1 \) when \( c_0 \in [0, 1] \) and \( c_1 \geq 0 \). Intuitively, we expect \( \alpha(SL) \) to be decreasing in \( SL \) (i.e., \( c_1 \geq 1 \)), since the lost-sales system behaves more like the backorder system for high service levels (see Lemma 2 for the connection between \( b_p \) and the backorder system).

The form in (8) is what we have referred to as a power approximation. Such approximations have been proposed earlier in the inventory literature by Ehrhardt [1979], Ehrhardt and Mosier [1984], and Schneider and Riguest [1990] in the context of computing near-optimal \((s, S)\) values for single-stage inventory systems with backordering and fixed ordering costs. To find the best values for the constants, we use least-square regression on the error terms,
which measures the distance from the proposed value of \(\alpha, \alpha(p/(p + \sum_i h_i))\), to the target interval \([\underline{\alpha}(h, \tau, p), \overline{\alpha}(h, \tau, p)]\), i.e.,

\[
\xi(h, \tau, p) = \begin{cases} 
\underline{\alpha}(h, \tau, p) - \alpha(p/(p + \sum_i h_i)) & \text{if } \alpha(p/(p + \sum_i h_i)) \leq \underline{\alpha}(h, \tau, p), \\
\alpha(p/(p + \sum_i h_i)) - \overline{\alpha}(h, \tau, p) & \text{if } \alpha(p/(p + \sum_i h_i)) \geq \overline{\alpha}(h, \tau, p), \\
0 & \text{otherwise.}
\end{cases}
\] (9)

As a final result of our regression we obtain \(c_0 = 0.441\) and \(c_1 = 17.236\). Since \(c_0 \in [0, 1]\) and \(c_1 \geq 0\), it follows that \(0 \leq \alpha(SL) \leq 1\). Now that we have a specific choice for \(\alpha(SL)\), which in turn yields a specific choice for \(\theta(p)\), we proceed to evaluate the cost-effectiveness of this approximation procedure in the next section.

7. Numerical Results

The goal of this section is threefold. First, we illustrate the asymptotic optimality of Theorem 1 for a wide range of parameter values. Second, we compare the performance of base-stock policies to the optimal policy. Third, the performance of the base-stock policy that is based on the power approximation procedure is investigated.

Besides the test bed described in Table 1, we also consider parameter settings based on three echelon levels and the Negative Binomial distribution for demand. As a result, we are able to test the performance of the power approximation procedure for different instances than the one used to construct the power approximation. When there are three echelon levels (i.e., \(M = 3\)), we set the lead time vector \(\tau = (1, 2, 1)\) such that the total lead time equals four review periods (similar to Table 1). When demand follows a negative binomial distribution, the value of mean demand is either 4 or 8 with a variance-to-mean (VTM) ratio of 2 or 4 (i.e., a multiple of the VTM ratio for the Poisson demand process). The parameter values for the unit penalty costs and unit holding costs are the same as described in Table 1. All combinations of these parameter settings yield an extra 72 instances for the setting with Poisson demand and three echelon levels, and 192 instances for the setting with Negative Binomial demands and 2 and 4 echelon levels.
<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$p$</th>
<th>$SL$</th>
<th>$C_{L^*}$</th>
<th>$b_p$</th>
<th>$S(b_p)$</th>
<th>$C_{L^*,S(b_p)}$</th>
<th>$S^*$</th>
<th>$CIS^*$</th>
<th>$\theta(SL)$</th>
<th>$S(\theta(SL))$</th>
<th>$CIS(\theta(SL))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>9</td>
<td>90.0%</td>
<td>9.73</td>
<td>12</td>
<td>18.27</td>
<td>1.15</td>
<td>16.23</td>
<td>3.01%</td>
<td>2.35</td>
<td>16.23</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td>19</td>
<td>95.0%</td>
<td>11.59</td>
<td>22</td>
<td>19.29</td>
<td>1.10</td>
<td>18.26</td>
<td>2.09%</td>
<td>6.95</td>
<td>18.26</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td>39</td>
<td>97.5%</td>
<td>13.25</td>
<td>42</td>
<td>20.30</td>
<td>1.05</td>
<td>19.28</td>
<td>1.17%</td>
<td>17.41</td>
<td>19.28</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>98.3%</td>
<td>14.13</td>
<td>62</td>
<td>21.31</td>
<td>1.04</td>
<td>20.29</td>
<td>0.87%</td>
<td>28.24</td>
<td>20.29</td>
<td>0.00%</td>
</tr>
<tr>
<td>4</td>
<td>79</td>
<td>98.8%</td>
<td>14.72</td>
<td>82</td>
<td>22.31</td>
<td>1.03</td>
<td>21.30</td>
<td>0.90%</td>
<td>39.17</td>
<td>20.30</td>
<td>0.11%</td>
</tr>
<tr>
<td>4</td>
<td>99</td>
<td>99.0%</td>
<td>15.17</td>
<td>102</td>
<td>22.32</td>
<td>1.03</td>
<td>21.30</td>
<td>0.64%</td>
<td>50.16</td>
<td>21.30</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>90.0%</td>
<td>16.07</td>
<td>12</td>
<td>33.50</td>
<td>1.13</td>
<td>29.44</td>
<td>2.61%</td>
<td>2.35</td>
<td>29.44</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>19</td>
<td>95.0%</td>
<td>18.64</td>
<td>22</td>
<td>34.52</td>
<td>1.07</td>
<td>32.48</td>
<td>1.65%</td>
<td>6.95</td>
<td>32.48</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>39</td>
<td>97.5%</td>
<td>20.89</td>
<td>42</td>
<td>36.54</td>
<td>1.05</td>
<td>34.51</td>
<td>1.01%</td>
<td>17.41</td>
<td>34.51</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>59</td>
<td>98.3%</td>
<td>22.08</td>
<td>62</td>
<td>37.55</td>
<td>1.04</td>
<td>35.53</td>
<td>0.87%</td>
<td>28.24</td>
<td>35.53</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>79</td>
<td>98.8%</td>
<td>22.87</td>
<td>82</td>
<td>37.56</td>
<td>1.03</td>
<td>36.54</td>
<td>0.75%</td>
<td>39.17</td>
<td>36.54</td>
<td>0.00%</td>
</tr>
<tr>
<td>8</td>
<td>99</td>
<td>99.0%</td>
<td>23.48</td>
<td>102</td>
<td>38.56</td>
<td>1.02</td>
<td>36.54</td>
<td>0.64%</td>
<td>50.16</td>
<td>36.54</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>9</td>
<td>90.0%</td>
<td>27.37</td>
<td>12</td>
<td>61.94</td>
<td>1.11</td>
<td>55.86</td>
<td>2.21%</td>
<td>2.35</td>
<td>55.85</td>
<td>0.08%</td>
</tr>
<tr>
<td>16</td>
<td>19</td>
<td>95.0%</td>
<td>30.93</td>
<td>22</td>
<td>62.97</td>
<td>1.06</td>
<td>59.91</td>
<td>1.45%</td>
<td>6.95</td>
<td>59.91</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>39</td>
<td>97.5%</td>
<td>34.02</td>
<td>42</td>
<td>64.100</td>
<td>1.04</td>
<td>62.96</td>
<td>0.94%</td>
<td>17.41</td>
<td>62.96</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>59</td>
<td>98.3%</td>
<td>35.65</td>
<td>62</td>
<td>65.101</td>
<td>1.03</td>
<td>64.98</td>
<td>0.76%</td>
<td>28.24</td>
<td>63.98</td>
<td>0.01%</td>
</tr>
<tr>
<td>16</td>
<td>79</td>
<td>98.8%</td>
<td>36.74</td>
<td>82</td>
<td>66.102</td>
<td>1.02</td>
<td>64.99</td>
<td>0.64%</td>
<td>39.17</td>
<td>64.99</td>
<td>0.00%</td>
</tr>
<tr>
<td>16</td>
<td>99</td>
<td>99.0%</td>
<td>37.56</td>
<td>102</td>
<td>67.103</td>
<td>1.02</td>
<td>65.100</td>
<td>0.52%</td>
<td>50.16</td>
<td>65.100</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Table 2: The results when demand follows a Poisson process and the system has 2 echelon levels with $h_i = 0.5$ for $i = 1, 2$.

### 7.1 Asymptotic Optimality

We first examine the fraction expressed in Theorem 1. Table 2 illustrates the performance for 18 base case instances; The parameter settings for these instances are $M = 2$ echelon levels, holding cost $h_i = 0.5$, penalty cost $p \in \{9, 19, 39, 59, 79, 99\}$ and a Poisson demand process with mean $\lambda$ equal to 4, 8 and 16. Based on these results, it is clear that the ratio $C_{L, S(b_p)}/C_{L^*}$ is close to 1 for moderately large values of $p$. When we consider all instances, Figure 1 and Figure 2 present the convergence over different values of the implied service level $SL$ when demand follows a Poisson and negative binomial distribution, respectively. We also observe from our experimental data that, for a fixed total lead time, the ratio tends to converge faster to 1 when more echelon levels are involved, the demand is higher or the VTM ratio is lower.
7.2 Base-stock policies

For each parameter setting we compute the optimal and best base-stock policy, and compare the expected total costs per period, which are denoted by $C^{L,*}$ and $C^{L,S,*}$, respectively. We are interested in the performance of the best base-stock policies compared to the optimal policies. Therefore, we use the relative cost increase between the two policies as our performance measure:

$$CI^{S,*} = 100\% \times \frac{C^{L,S,*} - C^{L,*}}{C^{L,*}}.$$  

Table 2 illustrates the performance for the base case instances. Based on these results, we conclude that in general $CI^{S,*}$ decreases when the average demand increases. However, the performance of base-stock policies mainly depends on the number of echelon levels as well as on the implied service level $SL$. Therefore, we depict $CI^{S,*}$ as a function of $SL$ over all instances in Figure 3 when demand follows a Poisson process. As a result, we conclude that base-stock policies perform well in multi-echelon settings. The cost increase for using base-stock policies compared to the optimal policy is 1.5%, 1.4% and 1.2% on average for inventory systems with 2, 3 and 4 echelon levels, respectively. Similar conclusions can be
found when demand is negative binomial, with an average cost increase of 2.0% and 1.7% for inventory systems with 2 and 4 echelon levels, respectively (see Table 4). The overall average cost increase is 1.6%. We also note that the maximum cost increase for the Poisson case is 4.5%, 3.5% and 2.6% for 2, 3 and 4 echelons, respectively; in the negative binomial case, the maximum cost increase is 4.7% and 3.4% for 2 and 4 echelons, respectively.

It is interesting to see that base-stock policies perform closer to optimal when more echelon levels are involved in the inventory system, or when the VTM ratio is lower. Furthermore, we notice in Figure 3 that $C^{S*}$ becomes zero when $SL$ decreases sufficiently enough. This is because the base-stock levels become zero when the penalty cost is relatively low, whereas the optimal policy prescribes never to order as well. As a result, we only consider instances with $SL \geq 75\%$ in the remainder of this section.

![Figure 3: The relative cost increase of the best base-stock policies over the optimal policy when demand follows a Poisson process.](image)

### 7.3 Power Approximation

We proceed with a thorough analysis of the performance of the power approximation. Let $C^{L,S(\theta(SL))}$ and $C^{L,S*}$ be the expected total cost per period when the inventory system is controlled using the power approximation and the best base-stock policy, respectively. Our measure of performance for a single item is the relative cost increase for using this approxi-
mation procedure compared to the best base-stock policy:

\[ CI^{S(\theta(SL))} = 100\% \times \frac{C^{L,S(\theta(SL))} - C^{L,S^*}}{C^{L,S^*}}. \]

Our results for inventory systems with parameter settings other than those used to derive the power approximation are presented in Table 3 and Table 4, i.e., Poisson demand with \( M = 3 \) and negative binomial demand with \( M = 2 \), respectively. Figure 4 and Figure 5 illustrate the average cost increase for the base-stock policy based on the power approximation compared to the best base-stock policy over all instances for various ranges of \( SL \), whereas a summary of these results are presented in Table 5. The results clearly show that the power approximation is highly accurate. In all instances with a \( SL \) between 75-85\%, it yields an expected cost increase within 0.8\% and within 0.4\% when \( SL \geq 85\% \). When we compare the performance of the instances used in the development of the power approximation (column 2 and 4 in Table 5) to the instances having a different parameter setting (column 3 and 5-8 in Table 5), the former set of instances tend to have lower \( CI^{S(\theta(SL))} \) values since they are used in the regression, but we find that the latter set of instances also have reasonably good \( CI^{S(\theta(SL))} \) values. The average \( CI^{S(\theta(SL))} \) value is 0.04\% for the former set, and 0.18\% for the latter set. We conclude that the power approximation is robust since it performs very accurate for a broader family of parameter values than those used in setting the coefficients for the power approximation.

8. Conclusions

In this paper, we study the performance of echelon order-up-to policies in multi-stage inventory systems with lost sales. We establish the asymptotic optimality of a large family of policies within this class as the service levels approach one. We also propose an easy-to-compute power approximation for the echelon order-up-to vector. We show numerically that the best base-stock policy and the policy prescribed by the power approximation provide excellent cost performance for a wide range of parameter settings when the implied service level is 75\% or more. More precisely, the cost of the best base-stock policy (power approximation) is, on an average, only 1.6 \% (0.2 \%) higher than the optimal cost (cost of the best base-stock policy).
<table>
<thead>
<tr>
<th>$h_i$</th>
<th>$p$</th>
<th>$SL$</th>
<th>$\lambda = 4$</th>
<th>$\lambda = 8$</th>
<th>$\lambda = 16$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$S^*$</td>
<td>$CIS^* S(\theta(SL))$</td>
<td>$CIS(\theta(SL))$</td>
</tr>
<tr>
<td>0.25</td>
<td>9</td>
<td>92.31%</td>
<td>12,20,24</td>
<td>2.09%</td>
<td>22,38,46</td>
</tr>
<tr>
<td>0.25</td>
<td>19</td>
<td>96.20%</td>
<td>14,23,27</td>
<td>1.25%</td>
<td>24,41,49</td>
</tr>
<tr>
<td>0.25</td>
<td>39</td>
<td>98.11%</td>
<td>15,25,29</td>
<td>0.75%</td>
<td>26,44,52</td>
</tr>
<tr>
<td>0.25</td>
<td>59</td>
<td>98.74%</td>
<td>16,26,30</td>
<td>0.59%</td>
<td>27,45,54</td>
</tr>
<tr>
<td>0.25</td>
<td>79</td>
<td>99.06%</td>
<td>16,26,31</td>
<td>0.51%</td>
<td>27,46,55</td>
</tr>
<tr>
<td>0.25</td>
<td>99</td>
<td>99.25%</td>
<td>16,27,31</td>
<td>0.41%</td>
<td>28,47,56</td>
</tr>
<tr>
<td>0.5</td>
<td>9</td>
<td>85.71%</td>
<td>11,18,21</td>
<td>2.82%</td>
<td>20,34,41</td>
</tr>
<tr>
<td>0.5</td>
<td>19</td>
<td>92.68%</td>
<td>13,21,24</td>
<td>2.14%</td>
<td>23,39,46</td>
</tr>
<tr>
<td>0.5</td>
<td>39</td>
<td>96.30%</td>
<td>14,23,27</td>
<td>1.18%</td>
<td>24,42,50</td>
</tr>
<tr>
<td>0.5</td>
<td>59</td>
<td>97.52%</td>
<td>15,24,28</td>
<td>0.86%</td>
<td>25,43,51</td>
</tr>
<tr>
<td>0.5</td>
<td>79</td>
<td>98.14%</td>
<td>15,25,29</td>
<td>0.73%</td>
<td>26,44,53</td>
</tr>
<tr>
<td>0.5</td>
<td>99</td>
<td>98.51%</td>
<td>15,25,30</td>
<td>0.70%</td>
<td>26,45,53</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>75.00%</td>
<td>9,14,16</td>
<td>3.46%</td>
<td>17,29,35</td>
</tr>
<tr>
<td>1</td>
<td>19</td>
<td>86.36%</td>
<td>11,18,21</td>
<td>2.78%</td>
<td>20,35,42</td>
</tr>
<tr>
<td>1</td>
<td>39</td>
<td>92.86%</td>
<td>13,21,25</td>
<td>2.11%</td>
<td>23,39,46</td>
</tr>
<tr>
<td>1</td>
<td>59</td>
<td>95.16%</td>
<td>13,22,26</td>
<td>1.45%</td>
<td>24,40,48</td>
</tr>
<tr>
<td>1</td>
<td>79</td>
<td>96.34%</td>
<td>14,23,27</td>
<td>1.15%</td>
<td>24,42,50</td>
</tr>
<tr>
<td>1</td>
<td>99</td>
<td>97.06%</td>
<td>14,23,27</td>
<td>1.11%</td>
<td>25,43,51</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>60.00%</td>
<td>6,6,7</td>
<td>0.63%</td>
<td>15,18,22</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>76.00%</td>
<td>9,14,17</td>
<td>3.37%</td>
<td>18,30,36</td>
</tr>
<tr>
<td>2</td>
<td>39</td>
<td>86.67%</td>
<td>11,18,21</td>
<td>2.77%</td>
<td>20,35,42</td>
</tr>
<tr>
<td>2</td>
<td>59</td>
<td>90.77%</td>
<td>12,20,23</td>
<td>2.20%</td>
<td>22,38,45</td>
</tr>
<tr>
<td>2</td>
<td>79</td>
<td>92.94%</td>
<td>13,21,25</td>
<td>2.05%</td>
<td>23,39,46</td>
</tr>
<tr>
<td>2</td>
<td>99</td>
<td>94.29%</td>
<td>13,21,25</td>
<td>1.58%</td>
<td>23,40,48</td>
</tr>
</tbody>
</table>

Table 3: The performance of the best base-stock policy and the power approximation when the demand is Poisson distributed and the system has 3 echelon levels.
Table 4: The performance of the best base-stock policy and the power approximation when the demand has a negative binomial distribution and the system has 2 echelon levels.
Figure 4: The relative cost increase of the power approximation over the best base-stock policy when the demand has a Poisson distribution.

Figure 5: The relative cost increase of the power approximation over the best base-stock policy when the demand has a negative binomial distribution.

Table 5: The average cost increase for the power approximation $CJ^S(\theta(SL))$ for different values of the implied service level $SL$. When no instance complies with the $SL$ interval, the entry is empty.
A. Proofs for Section 4.1

Proof of Lemma 2

Proof. In this proof, we fix a feasible policy for $L(h, \tau, \tilde{p})$ with $\tilde{p} = p + \sum_{m=2}^{M} h_m \cdot \sum_{j=1}^{m-1} \tau_j$, based on which we construct a feasible policy for $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$. We will obtain the required inequality by comparing the costs incurred by these two policies in their respective systems. We assume both systems start from the same state in period 1. This is without loss of generality since the optimal long run average cost does not depend on the starting state.

Under the given policy for $L(h, \tau, \tilde{p})$, let $q_{1,t}^L$ and $q_{2,t}^L$ denote the quantities ordered by echelons 1 and 2, respectively, in period $t$. Let $LOST_t$ denote the amount of lost sales incurred in $L(h, \tau, \tilde{p})$ in period $t$ under this policy.

We construct a policy for $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ by defining ordering quantities as follows. Let

$$q_{1,t}^B = q_{1,t}^L + LOST_{t-1-\tau_2}$$

$$and$$

$$q_{2,t}^B = q_{2,t}^L + LOST_{t-1}.$$ 

In any backordering system with time-invariant cost parameters, it is easy to prove that it is optimal to give units to customers as available without delay (in other words, use a first-come-first-served policy). Yet, we consider the following suboptimal policy of serving customers in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$. When a customer arrives, if there is a unit available, then it is given to her. Otherwise, this customer is moved to a backorder queue and a unit is ordered specifically for this customer from the external supplier in this period. This unit arrives $\tau_1 + \tau_2$ periods later and is then given to the customer.

The combination of this ordering policy and this service policy ensures that the number of customers that arrive in period $t$ and are not given a unit immediately in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ is the same as the amount of lost sales in $L(h, \tau, \tilde{p})$ in period $t$. Also, under this service policy, any backordered customer waits exactly $\tau_1 + \tau_2 + 1$ periods for her unit. Since the backorder cost parameter in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ is only $p/(\tau_1 + \tau_2 + 1)$, the total backorder costs accumulated in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ for each backordered demand is exactly $p$. An
additional $\sum_{m=2}^{M} h_m \cdot \sum_{j=1}^{m-1} \tau_j$ cost is incurred in holding costs due to the extra time that the backordered units are in transit in $B$ for this service policy. The sum of these costs is the same as the corresponding penalty cost per customer in $L(h, \tau, \tilde{p})$, which is $\tilde{p}$. Furthermore, it is easy to observe that the holding costs at both echelons (excluding the above-mentioned holding costs for "special ordered" units) are equal in $L(h, \tau, \tilde{p})$ and $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ in every period.

As a result, the total cost in $L(h, \tau, \tilde{p})$ is the same as the total cost incurred in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ under the above-mentioned ordering policy and service policy. Therefore, if the optimal service policy (that is, the first-come-first-served policy) is used instead, then the cost in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ would not increase. Thus, we have shown that for any feasible policy for $L(h, \tau, \tilde{p})$, we can construct a policy for $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ such that the cost in $B(h, \tau, p/(\tau_1 + \tau_2 + 1))$ is at most the cost in $L(h, \tau, \tilde{p})$ in every period. This implies the first inequality in the statement of the lemma.

The second inequality follows from the first inequality through a direct algebraic manipulation. 

Proof of Lemma 3

Proof. Suppose that $L(h, \tau, p)$ and $B(h, \tau, b_p)$ are both managed by the same echelon base-stock policy with an order-up-to vector $S = (S_1, S_2)$. We assume without loss of generality that both systems have the same initial state in the first period. Let $x_{l}^{C}$ and $x_{l}^{B}$ denote the echelon-1 on-hand inventory in $L(h, \tau, p)$ and the echelon-1 net-inventory in $B(h, \tau, b_p)$, respectively, at the beginning of period $t$ before observing demand. We denote the amount of lost sales in $L(h, \tau, p)$ in period $t$ by $LOST_{t}$ and the amount of backorders in $B(h, \tau, b_p)$ at the end of period $t$ by $BO_{t}$. For any $t \leq 0$, we let $D_{t} = 0$, $LOST_{t} = 0$ and $BO_{t} = 0$. Recall that $D_{[t_1, t_2]} = \sum_{t=t_1}^{t_2} D_{t}$ represents the cumulative demand incurred in the interval $[t_1, t_2]$. The quantities $LOST_{[t_1, t_2]}$ and $BO_{[t_1, t_2]}$ are similarly defined. Let $IP_{1,t}^{C}$ and $IP_{1,t}^{B}$ denote the echelon-1 inventory positions in period $t$ in $L(h, \tau, p)$ and $B(h, \tau, b_p)$, respectively, after ordering decisions have been made.
We first establish some basic properties related to the dynamics of $L$ and $B$. The equations below follow directly from the definitions of the quantities involved in them: for $t \geq \tau_1 + 1$,

$$x_t^L = IP_{1,t-\tau_1}^L - D[t - \tau_1, t - 1] + LOST[t - \tau_1, t - 1], \quad \text{and} \quad (10)$$

$$x_t^B = IP_{1,t-\tau_1}^B - D[t - \tau_1, t - 1]. \quad (11)$$

The relationships below follow from the use of echelon order-up-to policies with parameters $S_1$ and $S_2$ and the fact that the amount of lost sales is nonnegative. For all $t \geq \tau_1 + 1$,

$$IP_{1,t-\tau_1}^C = \min(S_1, S_2 - D[t - \tau_1 - \tau_2, t - \tau_1 - 1] + LOST[t - \tau_1 - \tau_2, t - \tau_1 - 1]) \quad (12)$$

$$\geq \min(S_1, S_2 - D[t - \tau_1 - \tau_2, t - \tau_1 - 1])$$

$$= IP_{1,t-\tau_1}^B. \quad (13)$$

Moreover, for all $t$,

$$x_t^B = \min(S_2 - D[t - \tau_1 - \tau_2, t - 1], S_1 - D[t - \tau_1, t - 1]) \quad (14)$$

The following claims will be useful in providing an upper bound on the costs incurred in $\mathcal{L}(h, \tau, p)$.

- **Claim 1**: For all $t$, $\text{LOST}_t \leq BO_t$. To prove this claim, we combine the inequality in (13) with the expressions for $x_t^L$ and $x_t^B$ given in (10) and (11), to get

$$x_t^L \geq IP_{1,t-\tau_1}^C - D[t - \tau_1, t - 1] \geq IP_{1,t-\tau_1}^B - D[t - \tau_1, t - 1] = x_t^B \quad \text{for } t \geq \tau_1 + 1.$$ 

It can similarly be shown that $x_t^C \geq x_t^B$ for $t \in \{1, 2, \ldots, \tau_1\}$. Therefore,

$$\text{LOST}_t = (D_t - x_t^C)^+ \leq (D_t - x_t^B)^+ = BO_t \quad \text{for all } t.$$ 

- **Claim 2**: For all $t$, $x_t^C \leq x_t^B + \text{LOST}[t - \tau_1 - \tau_2, t - 1]$.

To see this claim, note that equations (10) and (12) imply that for all $t \geq \tau_1 + 1$,

$$x_t^C = IP_{1,t-\tau_1}^C - D[t - \tau_1, t - 1] + \text{LOST}[t - \tau_1, t - 1]$$

$$= \min\{S_1, S_2 - D[t - \tau_1 - \tau_2, t - \tau_1 - 1] + \text{LOST}[t - \tau_1 - \tau_2, t - \tau_1 - 1]\}$$

$$- D[t - \tau_1, t - 1] + \text{LOST}[t - \tau_1, t - 1]$$

$$\leq \min\{S_1, S_2 - D[t - \tau_1 - \tau_2, t - \tau_1 - 1]\}$$

$$- D[t - \tau_1, t - 1] + \text{LOST}[t - \tau_1 - \tau_2, t - 1],$$
where the last inequality follows from the identify \( \min\{\alpha_1, \alpha_2 + \alpha_3\} \leq \min\{\alpha_1, \alpha_2\} + \alpha_3 \).

The inequality above can similarly be established for \( t \in \{1, 2, \ldots, \tau\} \). Now, the claim follows from the expression of \( x_t^B \) in (14).

- **Claim 3:** The echelon-2 inventory level in \( L(h, \tau, p) \) at the end of period \( t - \tau \) equals

\[
S_2 - D[t - \tau - \tau_2, t - \tau] + LOST[t - \tau - \tau_2, t - \tau] \quad \text{for all } t \geq \tau + 1 .
\]

This result follows from the fact that echelon-2 orders up to \( S_2 \) in each period.

We now use the above claims to derive an upper bound on \( C^{LS}(h, \tau, p) \). By Claim 3,

\[
C^{LS}(h, \tau, p) = h_1 \cdot \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E\left[(x_t^L - D_t)^+\right] + h_2 \cdot \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[S_2 - D[t - \tau - \tau_2, t - \tau] + LOST[t - \tau - \tau_2, t - \tau]] + p \cdot \limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[LOST_t] .
\]

Observe that

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[LOST[t - \tau - \tau_2, t - \tau]] = \limsup_{T \to \infty} \frac{1}{T} \left[ \sum_{t=1}^{T} (\tau_2 + 1) \cdot E[LOST_t] \right] \leq \limsup_{T \to \infty} \frac{1}{T} \left[ \sum_{t=1}^{T} (\tau_2 + 1) \cdot E[BO_t] \right] ,
\]

where the last inequality follows from Claim 1.
Also, similarly, from Claims 1 and 2,

\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ (x_t^B - D_t)^+ \right] 
\leq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ (x_t^B + LOST[t - \tau_1 - \tau_2, t - 1] - D_t)^+ \right] 
\leq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ (x_t^B - D_t)^+ + LOST[t - \tau_1 - \tau_2, t - 1] \right] 
\leq \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ (x_t^B - D_t)^+ + (\tau_1 + \tau_2) \cdot BO_t \right].
\]

Thus, applying (15) and (16) along with Claim 1,

\[
C_{L,S}(h, \tau, p) 
\leq h_1 \cdot \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ (x_t^B - D_t)^+ + (\tau_1 + \tau_2) \cdot BO_t \right] 
+ h_2 \cdot \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E \left[ S_2 - D[t - \tau_1 - \tau_2, t - \tau_1] + (\tau_2 + 1) \cdot BO_t \right] 
+ p \cdot \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[BO_t] 
= \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left[ h_1 \cdot E[(x_t^B - D_t)] + h_2 \cdot (S_2 - (\tau_2 + 1) \cdot E[D]) + (b_p + h_1 + h_2) \cdot E[BO_t] \right] 
= C_{B,S}(h, \tau, b_p),
\]

since \(b_p = p + h_1 \cdot (\tau_1 + \tau_2) + h_2 \cdot \tau_2\) by definition. This completes the proof.

\textbf{Proof of Lemma 5}

\textit{Proof.} Let us first consider the case in which \(\beta \geq 1\). We now have

\[
\beta \cdot b \geq b \quad \text{and} \quad \frac{C_{B,S}(h, \tau, \beta \cdot b)}{C_{B,S}(h, \tau, b)} \geq 1.
\]

Thus, the desired result can be shown by establishing that \(\lim_{b \to \infty} \left( \frac{C_{B,S}(h, \tau, \beta \cdot b)}{C_{B,S}(h, \tau, b)} \right) \leq 1\), which we proceed to do.

Let \(S_1(b)\) and \(S_2(b)\) denote the echelon-1 and echelon-2 order-up-to levels followed by an optimal policy for \(B(h, \tau, b)\). We compare the costs incurred in this system with the costs incurred by \(B(h, \tau, \beta \cdot b)\) while using the same policy.
The following properties of $S_1(b)$ and $S_2(b)$ are well known (see, for example, Shang and Song [2003]):

\[ S_1(b) = F_{\tau_1+1}^{-1} \left( \frac{b + h_2}{b + h_2 + h_1} \right) \quad \text{and} \quad (18) \]

\[ S_2(b) \geq S_2^*(b) \quad \text{where} \quad S_2^*(b) = F_{\tau_1+\tau_2+1}^{-1} \left( \frac{b}{b + h_2 + h_1} \right) , \quad (19) \]

where $F_\tau$ represents the cumulative distribution function of the $\tau$-fold convolution of the demand distribution for any positive integer $\tau$.

Let $BO_b$ denote the expected amount on backorder (or equivalently the long run average of this amount) at the end of a period when an echelon order-up-to policy with parameters $S_1(b)$ and $S_2(b)$ are used. We claim

\[ \lim_{b \to \infty} \frac{b \cdot BO_b}{C^{B,\ast}(h, \tau, b)} = 0 . \quad (20) \]

To prove this claim, note that the following formula for $BO_b$ is well known, and easy to derive from (14):

\[ BO_b = E \left[ \max \left\{ (D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+, (D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+ \right\} \right] . \]

Thus, we can bound $BO_b$ from above as follows:

\[ BO_b \leq E \left[ (D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+ + (D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+ \right] \]

\[ \leq E \left[ (D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+ + (D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+ \right] . \quad (21) \]

Meanwhile, we can bound $C^{B,\ast}(h, \tau, b)$ from below by

\[ h_1 \cdot (S_1(b) - E[D] \cdot (\tau_1 + 1)) \quad (22) \]

and also by

\[ h_2 \cdot (S_2^*(b) - E[D] \cdot (\tau_1 + \tau_2 + 1)) . \quad (23) \]

Both these lower bounds are strictly positive for large enough $b$ since $S_1(b)$ and $S_2^*(b)$ are increasing functions which approach infinity as $b$ increases because $D$ is an unbounded random variable by assumption. Combining the properties stated above, we get, for sufficiently
large \( b \),

\[
\frac{b \cdot BO_b}{C^{B,*}(h, \tau, b)} 
\leq \frac{b \cdot E[(D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b))^+] \cdot \frac{E[D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b)]}{h_1 \cdot (S_1(b) - E[D] \cdot (\tau_1 + 1))} + \frac{b \cdot E[(D[1, \tau_1 + \tau_2 + 1] - S_2(b))^+] \cdot \frac{E[D[1, \tau_1 + \tau_2 + 1] - S_2(b)]}{h_2 \cdot (S_2(b) - E[D] \cdot (\tau_1 + \tau_2 + 1))}.}
\]

Using the expressions of \( S_1(b) \) and \( S_2(b) \) in (18) and (19), the right-hand side can be bounded above with conditional expectations as follows:

\[
\frac{b \cdot h_1}{b + h_1 + h_2} \cdot \frac{E[D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - S_1(b) \cdot D[\tau_2 + 1, \tau_1 + \tau_2 + 1] > S_1(b)]}{h_1 \cdot (S_1(b) - E[D] \cdot (\tau_1 + 1))} + \frac{b \cdot (h_1 + h_2)}{b + h_1 + h_2} \cdot \frac{E[D[1, \tau_1 + \tau_2 + 1] - S_2(b) \cdot D[1, \tau_1 + \tau_2 + 1] > S_2(b)]}{h_2 \cdot (S_2(b) - E[D] \cdot (\tau_1 + \tau_2 + 1))}.
\]

As \( b \to \infty \), we know that \( S_1(b) \) and \( S_2(b) \) also approach \( \infty \). By the assumption in the statement of Theorem 1, \( E[D[1, \tau_1 + \tau_2 + 1] - d|D[1, \tau_1 + \tau_2 + 1] > d]/d \) and \( E[D[\tau_2 + 1, \tau_1 + \tau_2 + 1] - d|D[\tau_2 + 1, \tau_1 + \tau_2 + 1] > d]/d \) both approach zero as \( d \to \infty \). Thus, we complete the proof of the claim in (20).

Now, the cost incurred by \( B(h, \tau, \beta \cdot b) \), when managed by order-up-to levels \( S_1(b) \) and \( S_2(b) \), is exactly the cost of \( B(h, \tau, b) \) plus the additional backorder cost \( [(\beta - 1) \cdot b] \cdot BO_b \). Therefore, the optimal cost \( C^{B,*}(h, \tau, \beta \cdot b) \) is bounded above by \( C^{B,*}(h, \tau, b) \) plus \( [(\beta - 1) \cdot b] \cdot BO_b \), and we obtain

\[
1 \leq \frac{C^{B,*}(h, \tau, \beta \cdot b)}{C^{B,*}(h, \tau, b)} \leq 1 + (\beta - 1) \cdot \frac{b \cdot BO_b}{C^{B,*}(h, \tau, b)},
\]

where the first inequality was shown in (17).

The claim in (20) implies that the limit of the right-hand side of the above expression as \( b \to \infty \) is 1. This implies the required result for the case in which \( \beta \geq 1 \).

The proof of the result when \( \beta < 1 \) follows from the case discussed so far by considering the limiting ratio

\[
\lim_{b \to \infty} \left( \frac{C^{B,*}(h, \tau, b)}{C^{B,*}(h, \tau, \beta b)} \right) = \lim_{b \to \infty} \left( \frac{C^{B,*}(h, \tau, (1/\beta)\tilde{b})}{C^{B,*}(h, \tau, b)} \right),
\]

where \( \tilde{b} = \beta \cdot b \). Since \( 1/\beta > 1 \), we know from the first case that the limit on the right-hand side is 1. Thus, the limit on the left-hand side is 1 which implies the desired result. \( \square \)
References


