Working Paper No. 20-43

Understanding the performance of capped base-stock policies in lost-sales inventory models

Linwei Xin
University of Chicago Booth School of Business

Submitted to *Operations Research*
manuscript (Please, provide the manuscript number!)

Understanding the performance of capped base-stock policies in lost-sales inventory models

Linwei Xin
Booth School of Business, University of Chicago, Chicago, IL 60637
Linwei.Xin@chicagobooth.edu

Single-sourcing lost-sales inventory systems with lead times are notoriously difficult to optimize. In this work, we propose a new family of capped base-stock policies and provide a new perspective of constructing a practical hybrid policy combining two well-known heuristics: base-stock and constant-order policies. Each capped base-stock policy is associated with two parameters: a base-stock level and an order cap. We prove that for any fixed order cap, the capped base-stock policy converges exponentially fast in the base-stock level to a constant-order policy, providing a theoretical foundation for a phenomenon that a capped dual-index policy converges numerically to a Tailored Base-Surge policy recently observed in Sun and Van Mieghem (2019) in a different but related dual-sourcing inventory model. As a consequence, there exists a sequence of capped base-stock policies that are asymptotically optimal as the lead time grows. We also numerically demonstrate its superior performance in general (including small lead times) by comparing with other well-known heuristics.

Key words: inventory, lost-sales, lead time, capped base-stock policy, constant-order policy.

History: This version: May 15, 2020

1. Introduction

Single-sourcing lost-sales inventory systems with lead times have a long and rich history in the inventory literature and the earliest study dates back to Karlin and Scarf (1958). Due to the complex structure of the optimal policy and computational intractability, lost-sales inventory systems are challenging to optimize and great efforts in the past focused on constructing various heuristic policies. Recently considerable attention was attracted to study a so-called constant-order policy (e.g., Reiman 2004, Zipkin 2008, Goldberg et al. 2016, Xin and Goldberg 2016, Bu, Gong and Yao 2019). This policy places a constant order every period and is simple to implement. Surprisingly, it demonstrates good performance for large lead times and is proved to be asymptotically optimal as the lead time grows (Goldberg et al. 2016, Xin and Goldberg 2016). However, it generally does not perform well for small lead times. As commented in Xin and Goldberg (2016): "*if one could prove that the constant-order policy performs well for small to moderate lead times, this would open the door for the creation of practical hybrid algorithms, which solve large dynamic programs when the lead time is small, and transition to simpler algorithms for larger lead times*". A fundamental question of how to construct a practical hybrid policy that performs well for both large and small lead times, perhaps by incorporating features from other good-performing heuristics (for example, base-stock policy), remains open.

In this work, we propose a new family of so-called *capped base-stock* policies and provide a new perspective of constructing such a practical hybrid policy by combining two well-known heuristics: base-stock and constant-order policies. Specifically, each capped base-stock policy is associated with two parameters: a base-stock level S and an order cap r. This new policy can be regarded as a base-stock policy but subject to an additional order cap. Alternatively, it can be regarded as a constant-order policy but subject to a constrained inventory position. We prove that for any r, the capped base-stock policy $\pi_{S,r}$ converges exponentially fast in the base-stock level S to the constant-order policy π_r that places a constant order r each period, providing a theoretical foundation for a phenomenon that a capped dual-index policy converges numerically to a Tailored Base-Surge (TBS) policy recently observed in Sun and Van Mieghem (2019) in a different but related dual-sourcing inventory model (see more discussions in Section 1.1). As a consequence, there exists a sequence of capped base-stock policies that are asymptotically optimal as the lead time grows. We also numerically demonstrate its superior performance in general (including small lead times) by comparing with other well-known heuristics. Although this work does not directly study the performance of constant-order policies for small lead times, the question raised in Xin and Goldberg (2016), we try to answer this question in an indirect way, by finding a new family of capped base-stock policies that is nearly optimal for large lead times and numerically performs well for small lead times. The superior performance is primarily driven by the new added order cap that helps smooth orders and lower the inventory level by avoiding chasing surging demands in the past and placing large orders all at once.

The family of capped base-stock policies enjoys the flexibility of two parameters and has the following attractive "self-correction" property: the best capped base-stock policy essentially behaves as a base-stock policy when the lead time is small or the penalty cost is large (namely, the base-stock level plays a key rule) and naturally converges to a constant-order policy as the lead time increases (namely, the order cap becomes more critical). This property is desirable due to the asymptotic optimality of constant-order and base-stock policies in the large lead time and large penalty cost regimes respectively, exactly as commented in Sun, Wang and Zipkin (2016): "*Although these asymptotic results do not necessarily imply that the optimal policy converges to a constant-order policy or an order-up-to policy when the corresponding parameter grows large, it is still of interest to check whether well-performing heuristic policies exhibit such tendencies*". Because base-stock policies have been widely accepted and implemented in practice and it might be easier to convince firms to implement base-stock policies than to implement constant-order policies, introducing the family of capped base-stock policies opens a door to implement a hybrid policy with "base-stock flavor" and good performance guarantee even for large lead time settings (settings where traditional base-stock policies perform poorly). Moreover, the suppliers can also benefit from adapting such capped policies due to the order smoothing effect that eliminates "large" orders.

Although the family of capped base-stock policies demonstrates good performance, there is one caveat: the cost function $C(\pi_{S,r})$ does not necessarily enjoy the convexity in the parameters and we demonstrate the non-convexity of the cost function $C(\pi_{S,r})$ in the base-stock level S for both optimal and non-optimal r (see Figure 2). By contrast, the cost function of a base-stock policy π_S

is always convex in S (e.g., Huh et al. 2009). We also note that $C(\pi_{S,r})$ seems to be quasi-convex in S (i.e., first decreasing and then increasing) although it is not convex.

This work is also related to inventory models with capacities (either deterministic or random), which have been extensively studied in the literature (see Federgruen and Zipkin 1986a, Federgruen and Zipkin 1986b, Ciarallo, Akella and Morton 1994, and the research thereafter). Note that in these models, the capacity constraints (either deterministic or random) are exogenous and strictly imposed. By contrast, in our lost-sales model, there are no such (hard) capacity constraints, which are only imposed by the capped base-stock policies.

1.1. Capped policies in the literature

The idea of introducing an order cap in this paper is directly motivated by Sun and Van Mieghem (2019), which solve a robust rolling horizon dual-sourcing inventory model with general lead times. The (robust) optimal policy turns out to be a dual-index policy that constrains or "caps" the slow order, the so-called *capped dual-index* policy. In other words, the policy is characterized by three parameters: two base-stock levels (for the slow and fast suppliers respectively) and a cap on the slow order. The presence of the additional order cap is the key feature distinguishing the capped dual-index policy from other existing heuristics. It provides order-smoothing and constrains high future holding costs. The intuition and benefits of order-smoothing under this policy are nicely demonstrated in Sun and Van Mieghem (2019). In addition, their numerical study evaluates the performance of the capped dual-index policy in the canonical dual-sourcing setting with independent and identically distributed (i.i.d.) demands. The results imply superior performance of the capped dual-index policies compared with other heuristics well studied in the dual-sourcing literature. Furthermore, the numerical results in Sun and Van Mieghem (2019) also demonstrate an interesting phenomenon that the percentage of slow orders that is capped increases and eventually approaches 100% as the lead time difference grows. This finding establishes an important connection between the capped base-stock policy and the so-called Tailored Base-Surge (TBS) policy that places a constant-order from the slow supplier in each period and follows a base-stock rule at the fast supplier. It suggests that as the lead time grows, the capped dual-index policy converges to the TBS policy, which has been proved to be asymptotically optimal in Xin and Goldberg (2018). Although it provides some first evidence that the capped dual-index (DI-cap) policy converges to the TBS policy, Sun and Van Mieghem (2019) comment that "*general investigation and theoretical proofs are beyond the scope of this paper*" and conjecture that "*DI-cap policy is also asymptotically optimal in the stochastic setting as* $L \to \infty$". Besides outperforming several well studied heuristics, the capped dual-index policy can even beat the Asynchronous Advantage Actor-Critic (A3C) algorithm, one of the most popular recent deep reinforcement learning algorithms use deep neural networks to approximate value/policy functions of MDPs (see Gijsbrechts et. al 2018).

As a first step to theoretically understand the performance of such capped policies, in this work we investigate capped base-stock policies in the single-sourcing lost-sales model, because a dual-sourcing inventory model can be regarded as a generalization of a single-sourcing lost-sales inventory model (Sheopuri, Janakiraman and Seshadri 2010). The family of capped base-stock policies is analogous to the family of capped dual-index policies in the dual-sourcing model. We prove the convergence to the constant-order policy, which is analogous to the TBS policy in the dual-sourcing model.

The high-level idea that a good policy should have orders to be *smoother* than the filled demands was raised as early as in Zipkin (2000) (see Section 9.6.5), and similar restricted base-stock and modified base-stock policies were studied in related inventory models (see Bijvank and Johansen 2012, Johansen 2013). We also note that the capped base-stock policy is a special case of the so-called vector base-stock policy named by Zipkin (2008), although the idea of vector base-stock policies was proposed even earlier in the literature. As the family of vector base-stock policies is quite broad, it is computationally expensive to compute the best one. The existing research has been focusing on a special case: the better vector base-stock policy named by Zipkin (2008), which essentially reduces the search from a vector to a single parameter. The capped base-stock policy seems to inherit most of the benefit of the vector base-stock policy and our numerical study will demonstrate that it outperforms the better vector base-stock policy.

When this paper was under the second-round review, we became aware of the unpublished manuscript Johansen and Thorstenson (2008), which was unavailable online (we thank Paul Zipkin for pointing out the relevance and Søren Glud Johansen for sharing a copy of this unpublished manuscript). Johansen and Thorstenson (2008) essentially studied the same policy (called restricted base-stock policy) in the same lost-sales model and demonstrated its good performance. The steps ahead (which we believe are important) in what follows as compared to Johansen and Thorstenson (2008) are: (1) we explain why it performs so well and when it brings the most benefits by building the connection to the existing good-performing base-stock and constant-order policies; (2) we theoretically prove its convergence to the constant-order policy, which implies its asymptotic optimality.

1.2. Outline of paper

The rest of the paper is organized as follows. We describe the model in Section 2, review the family of constant-order policies in Section 2.1, and introduce the family of capped base-stock policies in Section 2.2. We state the main theoretic results in Section 2.3 and provide proofs in Section 3. We conduct a concise (but targeted) numerical study and demonstrate (perhaps surprisingly) superior performance of proposed capped base-stock policies in Section 4. We also explain why the proposed policies perform so well in general and when the improvement can be significant compared with existing heuristics. Finally, we summarize our main contributions and propose several interesting directions for future research in Section 5.

2. Model description and problem statement

In this section, we define our lost-sales inventory model with lead times. Let $\{D_t\}_{t\geq 1}$ be a sequence of i.i.d. demands. Let L be the lead time and h, p be the per-unit holding and lost-sales costs respectively. In addition, let I_t denote the on-hand inventory and $\mathbf{x}_t = (x_{1,t}, \ldots, x_{L,t})$ denote the inventory pipeline vector at the beginning of time period t, where $x_{i,t}$ is the order placed in period $i+t-1-L$ and to be received in period $i+t-1$. Then the sequence of events in period t is as follows:

- a new amount of inventory $x_{1,t}$ is delivered and added to the on-hand inventory;
- a new order q_t is placed;
- the demand D_t is realized;
- costs for period t are incurred, and the on-hand inventory and pipeline vector are updated.

In particular, the on-hand inventory is updated according to $I_{t+1} = \max\{I_t + x_{1,t} - D_t,\ 0\}$, and the pipeline vector is updated such that $x_{1,t}$ is removed, $x_{i,t+1}$ is set equal to $x_{i+1,t}$ for $i \in [1, L-1]$, and $x_{L,t+1}$ is set equal to q_t (the new order placed). Note that q_t cannot depend on the realizations of future demand. We call such policies admissible policies, and denote the family of all admissible policies by Π. We assume that the inventory is initially empty, i.e., zero on-hand inventory $I_1 = 0$ and empty pipeline vector $\mathbf{x}_1 = \mathbf{0}$. Define C_t to be the sum of the holding and lost-sales costs in period t:

$$C_t \triangleq\ h\left(I_t + x_{1,t} - D_t\right)^+ + p\left(I_t + x_{1,t} - D_t\right)^-,$$

where $x^+ \triangleq \max\{x, 0\}$ and $x^- \triangleq \max\{-x, 0\}$. Let $C(\pi)$ to be the long-run average cost incurred by a policy π:

$$C(\pi) \triangleq\ \limsup_{T\to\infty} \frac{\sum_{t=1}^{T} \mathbb{E}\left[C_t\right]}{T}.$$

Then the corresponding lost-sales inventory problem is given by

$$\mathrm{OPT}(L) \triangleq\ \inf_{\pi\in\Pi} C(\pi). \tag{1}$$

Here we explicitly express the dependency of the optimal cost OPT on the lead time L, although OPT depends on other parameters as well.

2.1. Constant-order policy

In this section, we provide a brief review of the family of constant-order policies, which are studied earlier, e.g., in Reiman (2004), Zipkin (2008), Goldberg et al. (2016), Xin and Goldberg (2016), and Bu, Gong and Yao (2019). Each constant-order policy π_r places a constant-order r in every period, where $r \in [0, \mathbb{E}[D])$. It is well-known that its steady-state on-hand inventory (denoted as

I_r^∞) has the same distribution as the steady-state waiting time in the single-server queue with inter-arrival distribution D and constant processing time r. Specifically, $I_r^\infty \sim \max_{k\geq 0}\left\{\sum_{i=1}^k (r-D_i)\right\}$, where $X\sim Y$ denote equivalence in distribution between X and Y. Then the long-run average cost incurred by π_r has the following simple expression:

$$C(\pi_r) = h\mathbb{E}\left[I_r^\infty\right] + p\left(\mathbb{E}\left[D\right] - r\right), \tag{2}$$

where the second term $\mathbb{E}[D]-r$ captures the average amount of lost-sales, which equals the difference of the mean demand and mean supply. It is proved in Xin and Goldberg (2016) that the best constant-order policy converges to the optimal policy exponentially fast in the lead time (to be reviewed in Proposition 1). In addition, the best constant-order quantity can be efficiently computed thanks to the convexity of $C(\pi_r)$ in r (e.g., see Xin and Goldberg 2016).

2.2. Capped base-stock policy

In this section, we formally define the family of capped base-stock policies. Each capped base-stock policy $\pi_{S,r}$ is associated with two parameters (S,r), where S is the base-stock level and r is the order cap. In period t, $\pi_{S,r}$ places an order q_t to bring the inventory position (i.e., the sum of the on-hand inventory and pipeline vector) up to S but constrained by the order capacity r. Namely,

$$q_t = \min\left\{\left(S - I_t - \sum_{i=1}^{L} x_{i,t}\right)^+, \ r\right\}.$$

Note that $\pi_{S,r}$ is reduced to a base-stock policy π_S when r is sufficiently large (say $r\geq S$) so we can assume that $r\leq S$ without loss of generality.

As already described in the introduction section, the purpose of intentionally adding an order cap is to prevent the decision maker from chasing surging demands in the past and placing large orders all at once. For example, if the lead time L is large and the inventory system is in a position that there is no on-hand inventory and the inventory pipeline is empty, then following a base-stock policy a large order will be placed (due to the large lead time), which will take a long time to clear and cause high holding costs. Note that there is no restriction on the order cap r and in principle it can be greater than the mean demand. Indeed, based on the numerical results in Sections 4, the optimal order cap is usually greater than or equal to $\mathbb{E}[D]$ as long as the lead time is not too large. However, in Sections 2.3 and 3, we will mainly focus on the case that $r<\mathbb{E}[D]$ as we will compare $\pi_{S,r}$ with a constant-order policy π_r.

2.3. Main theoretic results

In this section, we theoretically verify the phenomenon empirically observed in Sun and Van Mieghem (2019) in a different but related dual-sourcing inventory model. Define $\phi_r(\theta)\triangleq\mathbb{E}\left[e^{\theta(r-D)}\right]$ for $\theta>0$.

THEOREM 1. *For each $r\in[0,\mathbb{E}[D])$, $S>(L+1)r$ and $\theta>0$ satisfying $\phi_r(\theta)<1$,*

$$C(\pi_{S,r})-C(\pi_r)\leq\ rpe^{-\theta(S-(L+1)r)}\frac{\phi_r(\theta)}{1-\phi_r(\theta)}. \tag{3}$$

The convergence of $C(\pi_{S,r})$ to $C(\pi_r)$ as $S\to\infty$ is intuitive and perhaps not surprising, because the chance that an order gets capped grows with S such that the policy $\pi_{S,r}$ eventually behaves as a constant-order policy. However, Theorem 1 shows that the gap between $C(\pi_{S,r})$ and $C(\pi_r)$ decays to zero exponentially fast in S with the rate $e^{-\theta}$, where θ satisfies $\phi_r(\theta)<1$ (such a θ alway exists due to the fact that $r<\mathbb{E}[D]$). Briefly speaking, this exponential rate of convergence corresponds to the rate of exponential decay of the probability of the waiting time in a related single-server queue (with inter-arrival distribution D and constant processing time r) being above a certain threshold x: $\mathbb{P}(I_r^\infty\geq x)$, which is approximately $e^{-x\theta_r^{max}}$ for large x from the theory of large deviations, where $\theta_r^{max}\triangleq\sup\{\theta>0:\phi_r(\theta)<1\}$. We will formally prove it in Lemma 3 and refer the interested reader to Ganesh, Metters and Wischik (2004) for more discussions on the applications of large deviations theory to queueing problems.

EXAMPLE 1 (CASE OF EXPONENTIALLY DISTRIBUTED DEMAND). Suppose D is exponentially distributed with rate λ. Then $\phi_r(\theta)=\frac{\lambda}{\lambda+\theta}e^{r\theta}$ for all $0\leq r<\lambda^{-1}$ and $\theta>0$. There exists a unique positive number θ_r^{max} to the equation $\phi_r(\theta)=1$. Note that θ_r^{max} is non-increasing in r such that

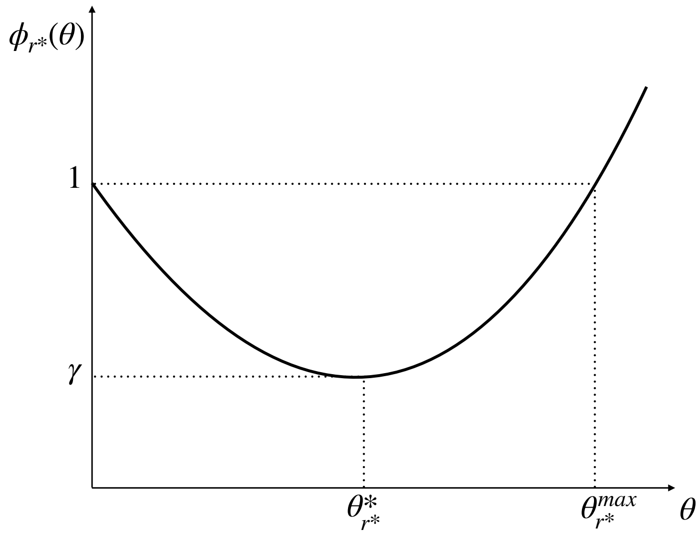

Figure 1 This graph illustrates the difference between the convergence rate γ in Proposition 1 and the convergence rate $e^{-\theta}$ (assuming $r = r^*$) in Theorem 1. γ corresponds to the convergence rate of the best constant-order policy to optimality (in L) and is achieved by $\theta^*_{r^*}$ that minimizes $\phi_{r^*}(\theta)$ over $[0,\infty)$. By contrast, $e^{-\theta}$ corresponds to the convergence rate of the capped base-stock policy π_{S,r^*} to the best constant-order policy π_{r^*} (in S) and θ can be arbitrary as long as $\theta < \theta^{max}_{r^*}$. The higher θ is, the faster the convergence is, but a higher θ may also lead to a larger prefactor in the bound (3).

the higher r is (interpreted as having a slower server in the related single-server queue), the slower the decay of the probability of the waiting time being above a certain threshold to zero.

Recall the following exponential convergence in Xin and Goldberg (2016). Let $r^* \in \arg\min_{r\in[0,\mathbb{E}[D])} C(\pi_r)$ be the best constant-order quantity.

Proposition 1 **(Theorem 1 of Xin and Goldberg 2016)**. *For all* $L \geq 1$, $\frac{C(\pi_{r^*})}{OPT(L)} \leq 1 + C_1\gamma^{L+1}$, *where* $\gamma \triangleq \min_{\theta\geq 0} \phi_{r^*}(\theta) < 1$ *and* $C_1 > 0$.

Note that the prefactor C_1 in Proposition 1 is explicitly given in Xin and Goldberg (2016). The exponential convergence rate γ relates to the expected waiting time in a single-server queue with inter-arrival distribution D and constant processing time r^* to its steady-state value. We illustrate the difference between γ and the convergence rate $e^{-\theta}$ (assuming $r = r^*$) in Theorem 1 in Figure 1.

As an immediate consequence of Theorem 1 and Proposition 1, we conclude that the family of capped base-stock policies is asymptotically optimal as the lead time grows.

COROLLARY 1. *There exists a sequence of* $\{(S_L, r_L)\}_{L\geq 1}$ *such that* $\lim_{L\to\infty} \frac{C(\pi_{S_L,r_L})}{OPT(L)} = 1$.

3. Proof of Theorem 1

In this section, we prove Theorem 1. We break the proof into two steps: (1) we derive an upper bound of $C(\pi_{S,r})$; (2) we use large deviations to analyze the waiting time of the related single-server queue and then bound the gap between the upper bound and $C(\pi_r)$.

Step 1: upper bound of $C(\pi_{S,r})$.

Under each capped base-stock policy $\pi_{S,r}$, the on-hand inventory and pipeline vector define a Markov chain and $C(\pi_{S,r})$ depends on the long-run behavior and steady-state distribution of the induced Markov chain. The following lemma establishes the existence of such a steady-state distribution. The idea is to use the empty inventory state (namely, empty pipeline vector and on-hand inventory) as regeneration points to yield a regenerative process, and then apply standard results in the theory of regenerative processes (e.g., Asmussen 2003) to prove the existence of the steady-state distribution. As the proof is quite similar to the argument used to prove the existence of the steady-state distribution under an optimal policy for the same lost-sales inventory problem (Theorem 2 in Xin and Goldberg 2016), we omit the details here.

LEMMA 1. *Under each policy* $\pi_{S,r}$*, there exists a steady-state distribution* $(\mathcal{I}, \mathfrak{q}_1, \ldots, \mathfrak{q}_L)$ *of the induced Markov chain, where* $\mathcal{I}$ *represents the inventory on-hand in the beginning of the period (before new order arrivals) and* $\mathfrak{q}_t$ *represents the order placed* $L+1-t$ *periods ago.*

Now we use Lemma 1 to obtain the following upper bound of $C(\pi_{S,r})$. Recall that I_r^∞ is the steady-state on-hand inventory under the constant-order policy π_r, defined in Section 2.1.

LEMMA 2. $C(\pi_{S,r}) \leq\ h\mathbb{E}[I_r^\infty] + p\,(\mathbb{E}[D] - r) + p\mathbb{E}\left[\left[(I_r^\infty + (L+1)r - S) \wedge r\right]^+\right]$.

Proof of Lemma 2. Note that the order placed by $\pi_{S,r}$ in each period is always no greater than r, such that the steady-state on-hand inventory under policy $\pi_{S,r}$ is bounded from above by I_r^∞. Hence with probability one (w.p.1), we have

$$\mathcal{I} \leq I_r^\infty, \quad 0 \leq \mathfrak{q}_i \leq r\ \forall i \in [1, L]. \tag{4}$$

Let $\mathfrak{q}_{L+1}$ be the order placed based on the inventory state $(\mathcal{I}, \mathfrak{q}_1, \ldots, \mathfrak{q}_L)$. From the definition of $\pi_{S,r}$,

$$\mathfrak{q}_{L+1} = \min\left\{\left(S - \mathcal{I} - \sum_{i=1}^{L} \mathfrak{q}_i\right)^+, \quad r\right\},$$

which implies w.p.1,

$$\begin{aligned} -\mathfrak{q}_{L+1} &= \max\left\{\left(\mathcal{I} + \sum_{i=1}^{L} \mathfrak{q}_i - S\right) \bigwedge 0, \quad -r\right\} \\ &\leq \max\left\{(I_r^\infty + Lr - S) \wedge 0, \quad -r\right\}, \end{aligned} \tag{5}$$

where $a \wedge b \triangleq \min\{a, b\}$. Combining stationarity with (4), (5) and the fact that the term $\mathbb{E}[D - \mathfrak{q}_{L+1}]$ captures the amount of long-run average lost-sales, we have

$$\begin{aligned} C(\pi_{S,r}) &= h\mathbb{E}[\mathcal{I}] + p\mathbb{E}[D - \mathfrak{q}_{L+1}] \\ &= h\mathbb{E}[\mathcal{I}] + p(\mathbb{E}[D] - r) + p\mathbb{E}[r - \mathfrak{q}_{L+1}] \\ &\leq h\mathbb{E}[I_r^\infty] + p(\mathbb{E}[D] - r) + p\mathbb{E}\left[\left[(I_r^\infty + (L+1)r - S) \wedge r\right]^+\right], \end{aligned}$$

completing the proof. $\square$

Step 2: bound the gap between the upper bound and $C(\pi_r)$.

Now we compare the upper bound derived in Lemma 2 with $C(\pi_r)$ in (2):

$$\begin{aligned} C(\pi_{S,r}) - C(\pi_r) &\leq p\mathbb{E}\left[\left[(I_r^\infty + (L+1)r - S) \wedge r\right]^+\right] \\ &\leq rp\mathbb{P}\left(I_r^\infty + (L+1)r - S \geq 0\right). \end{aligned} \tag{6}$$

It remains to bound the above probability.

LEMMA 3. *For any $r \in [0, \mathbb{E}[D])$, $x > 0$ and $\theta > 0$ such that $\phi_r(\theta) < 1$, we have $\mathbb{P}(I_r^\infty \geq x) \leq e^{-\theta x}\frac{\phi_r(\theta)}{1-\phi_r(\theta)}$.*

Proof of Lemma 3. Note that

$$\begin{aligned} \mathbb{P}(I_r^\infty \geq x) = \mathbb{P}\left(\max_{k\geq 1}\left\{\sum_{i=1}^{k}(r - D_i)\right\} \geq x\right) &\leq \sum_{k\geq 1}\mathbb{P}\left(\sum_{i=1}^{k}(r - D_i) \geq x\right) \\ &= \sum_{k\geq 1}\mathbb{P}\left(e^{\theta\left(\sum_{i=1}^{k}(r-D_i)\right)} \geq e^{\theta x}\right) \\ &\leq e^{-\theta x}\sum_{k\geq 1}(\phi_r(\theta))^k = e^{-\theta x}\frac{\phi_r(\theta)}{1-\phi_r(\theta)}, \end{aligned}$$

where the second inequality comes from Markov's inequality. The proof is completed. $\square$

Finally, combining (6) with Lemma 3 completes the proof of Theorem 1.

4. Numerical Experiments

In this section, we conduct a concise (but targeted) numerical study by using the same set of 32 problem instances as Zipkin (2008). Specifically, we normalize the holding cost at $h = 1$ and vary the penalty cost in the specified range, namely, $p \in \{4, 9, 19, 39\}$. We use two demand distributions: Poisson and Geometric, both with mean 5. The lead time is ranged from 1 to 4. We use the long-run average cost as the evaluating measure. We compare our capped base-stock policy with base-stock, constant-order and the so-called "myopic-2" policies. The myopic-2 policy uses a two-period horizon version of the myopic policy and is essentially within the framework of approximate dynamic programming. Note that Zipkin (2008) evaluates 8 heuristics in total including other well-known policies such as the dual-balancing policy, but myopic-2 is (arguably) the best performing heuristic. As such, we do not include other heuristics into our comparison. We also report the optimal value to Problem (1) (which is computed in Zipkin 2008) and percentages over the optimal of various heuristics.The results are summarized in Tables 1 and 2. In Table 3, we report the parameters of the best capped base-stock, base-stock, and constant-order policies. Although the parameters can be fractional in principle, for example, the best constant-order policy that is asymptotically optimal places fractional orders (e.g., see Xin and Goldberg 2016), our optimal parameters (i.e., base-stock levels, constant-order quantities and order caps) are searched only over integers in order to be consistent with numerical experiments conducted in Zipkin (2008) (among others). Finally, we use Matlab's solver "fmincon" to find the "best" capped base-stock policy, although the cost function is not necessarily convex in the parameters and there is no guarantee that the returned parameters are truly optimal (see Figure 2). Let us summarize our findings as follows.

First, Tables 1 and 2 demonstrate (perhaps surprisingly) superior performance of our capped base-stock policy: it is always within 1.5% of optimal (with an average of 0.71% optimality gap and 0.14% variance) over the 32 problem instances. The results in the tables also show that it dominates the base-stock and constant-order policies, although it is not surprising because both the base-stock and constant-order policies are special cases of capped base-stock policies. The improvement is the most significant when neither the constant-order nor base-stock performs well (e.g., $L=3,4$ and $p=4,9$). In particular, the capped base-stock policy significantly outperforms the constant-order policy when the lead time is small and/or the penalty cost is large (e.g., $L=1$ and $p=39$). It also significantly outperforms the base-stock policy when the lead time is large and the penalty cost is small (e.g., $L=4$ and $p=4$). By contrast, the improvement is not significant when either the constant-order or base-stock performs well (e.g., the ratio $\frac{L}{p}$ is very large or very small) due to their asymptotic optimality.

Second, from Table 3, the optimal order cap can be greater than or equal to the mean demand especially for small lead times or large penalty costs. By contrast, this is not allowed by the constant-order policy and the best constant-order quantity is always strictly less than the mean demand. In addition, the order cap of our capped base-stock policy is always greater than the constant-order quantity of the constant-order policy especially for instances with small lead times and large penalty costs, because these are the cases where the base-stock policy outperforms the constant-order policy such that the best capped base-stock policy tends to behave as a base-stock instead of a constant-order policy. Moreover, the base-stock level of our capped base-stock policy is always no smaller than that of the base-stock policy, partly to balance the effect of adding an order cap. Furthermore, the optimal order cap decreases as the lead time grows, because the best capped base-stock policy tends to behave as a constant-order policy such that its order cap eventually falls below the mean demand (which is five) as the lead time increases.

Third, from Tables 1 and 2, although our capped base-stock policy dominates the base-stock and constant-order policies, it is slightly worse than myopic-2 on average: myopic-2 has an average of 0.58% optimality gap, 0.35% variance, and 1.9% maximum, versus our capped base-stock policy having an average of 0.71% optimality gap, 0.14% variance, and 1.5% maximum. However, myopic-2 can be computationally expensive for large lead times and its performance seems to deteriorate as the lead time grows (e.g., Sun, Wang and Zipkin 2016). In addition, it is less interpretable than our capped base-stock policy. We also expand our test bed to incorporate a few instances with larger lead time values ($L = 6, 8, 10$) to test the robustness of the performance as the lead time grows, and report the results in Table 4. Note that in Table 4, CBS, BS, CO and M2 represent capped base-stock, base-stock, constant-order and myopic-2, respectively. Similar to numerical experiments conducted in other papers, we no longer report the optimal value to Problem (1) since computing the exact optimal policy becomes challenging when the lead time is beyond 4-period due to the curse of dimensionality. Table 4 shows that our capped base-stock policy dominates myopic-2 (as well as base-stock and constant-order policies) in all the instances, suggesting the robustness of the performance of the capped base-stock policy.

As final comments, we note that besides the heuristics tested in Zipkin (2008), there are recently developed heuristics using sophisticated approximate dynamic programming technique (e.g., Sun, Wang and Zipkin 2016), which demonstrate good performance. However, they are usually less interpretable than simple policies such as capped base-stock, base-stock and constant-order policies so we do not include them into the comparison here. In addition, our numerical experiments do not include instances with very large lead time (beyond 10) or very large penalty costs (beyond 39), because they are the two asymptotic regimes where the lost-sales problem becomes relatively easy to deal with and simple policies are asymptotically optimal. Furthermore, the constant-order policy performs better as the lead time grows due to its asymptotic optimality and can even beat myopic-2 in some cases (e.g., $L = 10$ and $p = 4$). However, it does not perform well for large p, partly because we only search over integer constant-order quantities. For demand distributions with mean five, the feasible set of constant-order quantities is only $\{0, 1, 2, 3, 4\}$ since the constant-order quantity has to be strictly less than the mean demand. However, if we allow fractional constant-order quantities, then the feasible set of constant-order quantities would be much larger and the cost would be reduced (sometimes significantly). For example, the cost would be significantly reduced from 40.27 to 18.3 when $p = 39$ and the demand distribution is Poisson.

5. Conclusion

Although there have been discussions on combining base-stock and constant-order policies to design a hybrid policy that performs well for any lead time values, how to implement this idea remains an open question. In this work, we effectively implemented this idea and proposed a new family of capped base-stock policies, incorporating the feature of base-stock policies into constant-order policies and significantly extending the applicability of constant-order policies especially for small lead times. We proved that for any fixed order cap, the capped base-stock policy converges exponentially fast in the base-stock level to a constant-order policy, providing a theoretical foundation for a phenomenon that a capped dual-index policy converges numerically to a TBS policy recently observed in Sun and Van Mieghem (2019) in a different but related dual-sourcing inventory model. As a consequence, we showed that there exists a sequence of capped base-stock policies that are asymptotically optimal as the lead time grows. We also numerically demonstrated its superior performance in general (including small lead times) by comparing with other well-known heuristics.

This work opens an important and promising new stream of research directions in classical inventory problems. In particular, one would expect that the idea of adding an order cap can be broadly extended to other inventory problems such as dual-sourcing inventory systems with general lead times. We expect similar asymptotic results to hold in the dual-sourcing setting (e.g., the convergence of the capped dual-index policy to the TBS policy as empirically observed in Sun and Van Mieghem 2019), but leave further investigation for future research. Moreover, this paper uses a "brute-force" search to find the best capped base-stock policy. It would be interesting to further explore the structure of the cost on the parameters (S, r) and develop more effective algorithms.

Acknowledgments

The author thanks Joachim Arts, Jinzhi Bu, Xin Chen, David A. Goldberg, Xiting Gong, Søren Glud Johansen, Cong Shi, Amy Ward, Hunan Zhang, Yuan Zhong, and Paul Zipkin for stimulating discussions. Research of the author was partly supported by the NSF award CMMI-1635160.

References

Asmussen, S. 2003. *Applied probability and queues*. Second Edition, Springer.

Bijvank, M., S. G. Johansen. 2012. Periodic review lost-sales inventory models with compound Poisson demand and constant lead times of any length. *European Journal of Operational Research*. **220** 106-114.

Bu, J., X. Gong, D. Yao. 2019. Constant-order Policies for Lost-sales Inventory Models with Random Supply Functions: Asymptotics and Heuristic. *Operations Research*. Forthcoming.

Ciarallo, F.W., R. Akella, T.E. Morton. 1994. A periodic review, production planning model with uncertain capacity and uncertain demand: optimality of extended myopic policies. *Management Science*. **40**(3) 320-332.

Federgruen, A., P. Zipkin. 1986a. An Inventory Model with Limited Production Capacity and Uncertain Demands I. The Average Cost Criterion. *Mathematics of Operations Research*. **11**(2) 193-207.

Federgruen, A., P. Zipkin. 1986b. An Inventory Model with Limited Production Capacity and Uncertain Demands II. The Discounted-Cost Criterion. *Mathematics of Operations Research*. **11**(2) 208-215.

Ganesh, A., N. O'Connell, D. Wischik. 2004. *Big Queues*. Springer, Berlin.

Gijsbrechts, J., R. N, Boute, J. A. Van Mieghem, D. Zhang. 2018. Can Deep Reinforcement Learning Improve Inventory Management? Performance and Implementation of Dual Sourcing-Mode Problems. Available at SSRN: https://ssrn.com/abstract=3302881.

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

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

Johansen, S.G. 2013. Modified base-stock policies for continuous-review, lost sales inventory models with poisson demand and a fixed lead time. *International Journal of Production Economics*. **143**(2) 379-384.

Johansen, S.G., A. Thorstenson. 2008. Pure and restricted base-stock policies for the lost-sales inventory system with periodic review and constant lead times. Technical Report, Department of Operations Research, Aarhus University.

Karlin, S., H. Scarf. 1958. Inventory models of the Arrow-Harris-Marschak type with time lag. K. Arrow, S. Karlin, H. Scarf (Eds.), *Studies in the Mathematical Theory of Inventory and Production*, Stanford University Press, Stanford, CA, 155-178.

Reiman, M.I. 2004. A new simple policy for a continuous review lost-sales inventory model. Unpublished manuscript.

Sheopuri, A., G. Janakiraman, S. Seshadri. 2010. New policies for the stochastic inventory control problem with two supply sources. *Operations Research.* **58**(3) 734-745.

Sun, J., J. A. Van Mieghem. 2019. Robust Dual Sourcing Inventory Management: Optimality of Capped Dual Index Policies and Smoothing. *Manufacturing & Service Operations Management.* **21**(4) 713-948.

Sun, P., K. Wang, P. Zipkin. 2016. Quadratic approximation of cost functions in lost sales and perishable inventory control problems. Working Paper (March 22, 2016).

Xin, L., D.A. Goldberg. 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., D.A. Goldberg. 2018. Asymptotic optimality of Tailored Base-Surge policies in dual-sourcing inventory systems. *Management Science.* **64**(1) 437-452.

Zipkin, P. 2000. *Foundations of inventory management.* McGraw-Hill, New York.

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

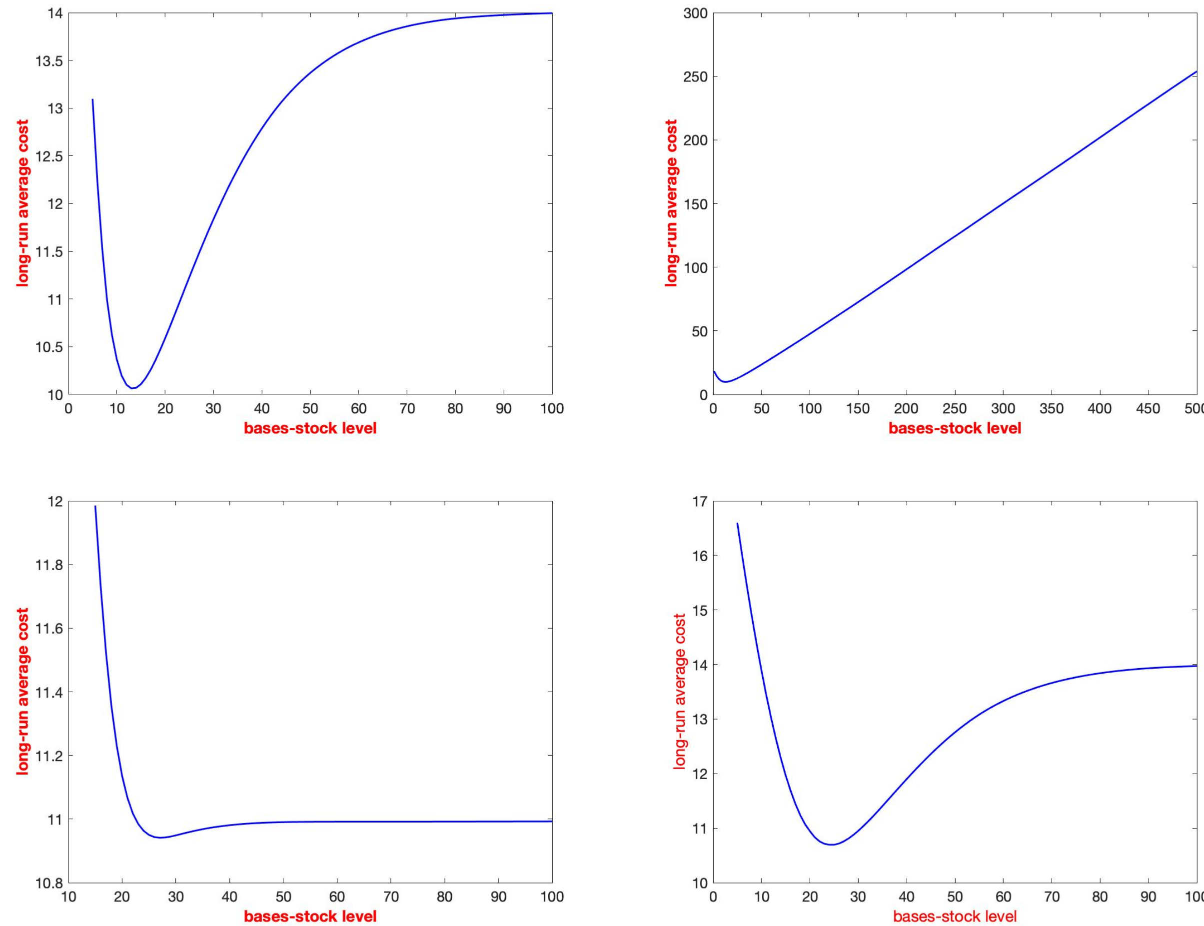

Figure 2 These graphs illustrate the non-convexity of $C(\pi_{S,r})$ in S for both optimal and non-optimal r. The demand distribution is Geometric with mean 5, $h=1$ and $p=4$. The first two graphs correspond to the cases with the same lead time $L=1$ but different order caps: non-optimal $r=4$ (Upper Left) and optimal $r=5$ (Upper Right). The last two graphs correspond to the cases with the same lead time $L=4$ but different order caps: non-optimal $r=3$ (Lower Left) and optimal $r=4$ (Lower Right). Note that the only exception of non-convexity is the upper right graph where the order cap is 5 (which equals the mean demand) such that the corresponding constant-order policy has an infinite long-run average cost. That is why the cost is monotone (after the turning point) and unbounded in S.

Long-run average cost and percentage over the optimal cost (Poisson)								
Policy	$L=1$		$L=2$		$L=3$		$L=4$	
(a) $p=4$								
Optimal	4.04		4.40		4.60		4.73	
Capped Base-stock	4.06	0.5%	4.41	**0.2%**	4.63	**0.7%**	4.80	**1.5%**
Base-stock	4.16	3.0%	4.64	5.5%	4.98	8.3%	5.20	9.9%
Constant-order	5.27	30.4%	5.27	19.8%	5.27	14.6%	5.27	11.4%
Myopic-2	4.04	**0.0%**	4.41	**0.2%**	4.64	0.9%	4.82	1.9%
(b) $p=9$								
Optimal	5.44		6.09		6.53		6.84	
Capped Base-stock	5.48	0.7%	6.12	0.5%	6.62	1.4%	6.91	**1.0%**
Base-stock	5.55	2.0%	6.32	3.8%	6.86	5.1%	7.27	6.3%
Constant-order	10.27	88.8%	10.27	68.6%	10.27	57.3%	10.27	50.1%
Myopic-2	5.44	**0.0%**	6.10	**0.2%**	6.57	**0.6%**	6.92	1.2%
(c) $p=19$								
Optimal	6.68		7.66		8.36		8.89	
Capped Base-stock	6.69	0.1%	7.72	0.8%	8.40	**0.5%**	8.95	**0.7%**
Base-stock	6.73	0.7%	7.84	2.3%	8.60	2.9%	9.23	3.8%
Constant-order	20.27	203%	20.27	165%	20.27	142%	20.27	128%
Myopic-2	6.68	**0.0%**	7.67	**0.1%**	8.40	**0.5%**	8.96	0.8%
(d) $p=39$								
Optimal	7.84		9.11		10.04		10.79	
Capped Base-stock	7.84	**0.0%**	9.14	0.3%	10.08	0.4%	10.88	0.8%
Base-stock	7.86	0.3%	9.19	0.9%	10.22	1.8%	11.06	2.5%
Constant-order	40.27	414%	40.27	342%	40.27	301%	40.27	273%
Myopic-2	7.84	**0.0%**	9.12	**0.1%**	10.07	**0.3%**	10.84	**0.5%**

Table 1 Numbers are marked in bold to highlight the best performing policy among the four heuristics.

Long-run average cost and percentage over the optimal cost (Geometric)								
Policy	$L=1$		$L=2$		$L=3$		$L=4$	
(a) $p=4$								
Optimal	9.82		10.24		10.47		10.61	
Capped Base-stock	9.87	0.5%	10.32	0.8%	10.51	**0.4%**	10.70	**0.8%**
Base-stock	10.04	2.2%	10.70	4.5%	11.13	6.3%	11.44	7.8%
Constant-order	11.00	12.0%	11.00	7.4%	11.00	5.1%	11.00	3.7%
Myopic-2	9.83	**0.1%**	10.29	**0.5%**	10.59	1.1%	10.80	1.8%
(b) $p=9$								
Optimal	14.51		15.50		16.14		16.58	
Capped Base-stock	14.58	0.5%	15.63	0.8%	16.27	**0.8%**	16.73	**0.9%**
Base-stock	14.73	1.5%	15.99	3.2%	16.87	4.5%	17.54	5.8%
Constant-order	19.00	30.9%	19.00	22.6%	19.00	17.7%	19.00	14.6%
Myopic-2	14.51	**0.0%**	15.57	**0.5%**	16.32	1.1%	16.89	1.9%
(c) $p=19$								
Optimal	19.22		20.89		22.06		22.95	
Capped Base-stock	19.32	0.5%	21.06	0.8%	22.27	1.0%	23.28	**1.4%**
Base-stock	19.40	0.9%	21.31	2.0%	22.73	3.0%	23.85	3.9%
Constant-order	29.00	50.9%	29.00	38.8%	29.00	31.5%	29.00	26.4%
Myopic-2	19.22	**0.0%**	20.95	**0.3%**	22.24	**0.8%**	23.29	1.5%
(d) $p=39$								
Optimal	23.87		26.21		27.96		29.36	
Capped Base-stock	24.00	0.5%	26.30	0.3%	28.28	1.1%	29.76	1.4%
Base-stock	24.00	0.5%	26.55	1.3%	28.51	2.0%	30.12	2.6%
Constant-order	49.00	105%	49.00	87.0%	49.00	75.3%	49.00	66.9%
Myopic-2	23.87	**0.0%**	26.26	**0.2%**	28.10	**0.5%**	29.64	**1.0%**

Table 2 Numbers are marked in bold to highlight the best performing policy among the four heuristics.

	Poisson with mean five				Geometric with mean five			
Policy	$L=1$	$L=2$	$L=3$	$L=4$	$L=1$	$L=2$	$L=3$	$L=4$
(a) $p=4$								
Capped Base-stock (S,r)	12,6	17,5	21,5	26,5	13,5	17,5	21,4	25,4
Base-stock (S)	12	16	20	25	12	15	18	21
Constant-order (r)	4	4	4	4	3	3	3	3
(b) $p=9$								
Capped Base-stock (S,r)	14,6	19,6	24,6	29,5	18,8	23,7	27,6	32,5
Base-stock (S)	13	19	23	28	17	22	25	30
Constant-order (r)	4	4	4	4	4	4	4	4
(c) $p=19$								
Capped Base-stock (S,r)	15,7	21,7	26,6	31,6	23,10	29,9	34,8	39,8
Base-stock (S)	15	20	26	31	22	28	32	37
Constant-order (r)	4	4	4	4	4	4	4	4
(d) $p=39$								
Capped Base-stock (S,r)	16,8	22,7	28,7	34,6	28,13	34,13	39,13	45,12
Base-stock (S)	16	22	28	33	27	34	39	45
Constant-order (r)	4	4	4	4	4	4	4	4

Table 3 Optimal parameters of heuristics.

		Poisson with mean five				Geometric with mean five			
L	p	CBS	BS	CO	M2	CBS	BS	CO	M2
6	4	**5.03**	5.51	5.27	5.05	**10.91**	11.86	11.00	11.08
6	9	**7.26**	7.90	10.27	7.43	**17.35**	18.53	19.00	17.75
6	19	**9.80**	10.20	20.27	9.85	**24.49**	25.54	29.00	24.93
6	39	**12.08**	12.38	40.27	12.11	**31.86**	32.69	49.00	32.12
8	4	**5.19**	5.72	5.27	5.20	**10.96**	12.12	11.00	11.27
8	9	**7.55**	8.32	10.27	7.77	**17.68**	19.18	19.00	18.39
8	19	**10.35**	10.90	20.27	10.53	**25.38**	26.81	29.00	26.21
8	39	**12.94**	13.39	40.27	13.09	**33.97**	34.47	49.00	34.12
10	4	**5.27**	5.86	**5.27**	5.31	**10.98**	12.31	11.00	11.40
10	9	**7.77**	8.63	10.27	8.08	**17.88**	19.68	19.00	18.89
10	19	**10.66**	11.48	20.27	11.09	**25.98**	27.82	29.00	27.27
10	39	**13.71**	14.24	40.27	13.93	**35.64**	36.25	49.00	35.82

Table 4 Long-run average costs. Numbers are marked in bold to highlight the best performing policy.

Brief Bio:

Linwei Xin is an assistant professor of Operations Management at the University of Chicago Booth School of Business. His primary research interest is supply chain and inventory management. His work has been recognized with several INFORMS paper competition awards, including the APS Best Publication Award (2019), First Place in the George E. Nicholson Student Paper Competition (2015), Second Place in the JFIG Paper Competition (2015), and a finalist in the MSOM Student Paper Competition (2014).