Asymptotic optimality of Tailored Base-Surge policies in dual-sourcing inventory systems

Linwei Xin
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

David A. Goldberg
School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Dual-sourcing inventory systems, in which one supplier is faster (i.e. express) and more costly, while the other is slower (i.e. regular) and cheaper, arise naturally in many real-world supply chains (cf. Allon and Van Mieghem (2010)). However, these systems are notoriously difficult to optimize due to the complex structure of the optimal solution and the curse of dimensionality. Indeed, as succinctly described in Veeraraghavan and Scheller-Wolf (2008) (and reiterated in Hua et al. (2014)),

“The dual-sourcing decision under general lead times has been a challenging problem for over 40 years, despite its frequency in practice.”

Recently, so-called Tailored Base-Surge (TBS) policies have been proposed as a heuristic for the dual-sourcing problem, and analyzed in Allon and Van Mieghem (2010) and Janakiraman, Seshadri and Sheopuri (2014). Under such a policy, a constant order is placed at the regular source in each period, while the order placed at the express source follows a simple order-up-to rule. Numerical experiments by several authors have suggested that such policies perform well as the lead time difference between the two sources grows large, which is exactly the setting in which the curse of dimensionality leads to the problem becoming intractable. However, providing a theoretical foundation for this phenomenon has remained a major open problem.

In this paper, we provide such a theoretical foundation by proving that a simple TBS policy is indeed asymptotically optimal as the lead time of the regular source grows large, with the lead time of the express source held fixed. Our main proof technique combines a steady-state approach, novel convexity and lower-bounding arguments, a certain interchange of limits result, and ideas from the theory of random walks and queues, significantly extending the methodology and applicability of a novel framework for analyzing inventory models with large lead times recently introduced in Goldberg et al. (2012) and Xin and Goldberg (2014) in the context of lost-sales models with positive lead times.

Key words: inventory, dual-sourcing, Tailored Base-Surge policy (TBS), lead time, asymptotic optimality, steady-state.
1. Introduction

Companies face the challenge of optimizing their sourcing strategies in a globalized world, and how to best utilize different sources effectively is a billion dollar industry. In practice, many companies (such as Caterpillar, cf. Rao, Scheller-Wolf and Tayur (2000)), often adopt dual-sourcing strategies for making such decisions. Under a dual-sourcing strategy, the companies usually purchase their materials from a regular supplier at a lower cost, but they are also able to obtain materials from an expedited supplier at a higher cost under emergency circumstances. For example, in the summer of 2003, Amazon used FedEx to deliver the new Harry Potter more promptly and maintained regular shipping via UPS (cf. Kelleher (2003), Veeraraghavan and Scheller-Wolf (2008)). Allon and Van Mieghem (2010) describes an example of a $10 billion high-tech U.S. company that has two suppliers, one in Mexico and one in China. The one in Mexico has shorter lead time but higher per-unit ordering cost; the one in China has longer lead time (5 to 10 times longer) but lower per-unit ordering cost. The company takes advantage of the dual-sourcing strategy to meet the demand more responsively (from Mexico) as well as less expensively (from China).

Although dual-sourcing is attractive, and very relevant to practice, optimizing a dual-sourcing inventory system is notoriously challenging. Such inventory systems have been studied now for over forty years, but the structure of the optimal policy remains poorly understood, with the exception of when the system is consecutive, i.e., the lead time difference between the two sources is exactly one. More specifically, the earliest studies of periodic review dual-sourcing inventory models include Barankin (1961), Daniel (1963), and Neuts (1964), which showed that base-stock (also known as order-up-to) policies are optimal when the lead times of the two sources are zero and one respectively. Fukuda (1964) extended the result to general lead time settings as long as the lead time difference remains one. Whittmore and Saunders (1977) showed that the optimal policy is no longer a simple base-stock policy when the lead time difference is beyond one and the
structure of the optimal policy can be quite complex. Furthermore, it is well known that a dual-sourcing inventory system can be regarded as a generalization of a lost-sales inventory system (cf. Sheopuri, Janakiraman and Seshadri (2010)). Indeed, the intractability of both the dual-sourcing and lost-sales inventory models has a common source - as the lead time grows, the state-space of the natural dynamic programming (DP) formulation grows exponentially, rendering such techniques impractical. This issue is typically referred to as the “curse of dimensionality” (cf. Karlin and Scarf (1958), Morton (1969), Zipkin (2008)), and we refer the reader to Goldberg et al. (2012) and Xin and Goldberg (2014) for a relevant discussion in the context of lost-sales inventory models.

There is a vast literature investigating periodic review dual-sourcing inventory models as well as their variants (cf. Minner (2003)), including: models with multiple suppliers (cf. Zhang (1996), Minner (2003), Feng et al. (2006)); models with two suppliers, one with higher variable costs but lower setup costs, and one with lower variable costs but higher setup costs (cf. Fox, Metters and Semple (2006)); models with a long-term contract supplier and a spot market (cf. Yi and Scheller-Wolf (2003), Chen, Xue and Yang (2013)); models for which the unmet demand must be satisfied from the expediting source (cf. Huggins and Olsen (2010)); models with expediting and advance demand information (cf. Angelus and Özer (2015)); models allowing emergency orders within the regular review period (cf. Targaras and Vlachos (2001)); models with correlated demand (cf. Sethi, Yan and Zhang (2003)); models considering capacity cost and flexibility for sourcing decisions (cf. Boute and Van Mieghem (2014)); multi-echelon models with expediting (cf. Lawson and Porteus (2000)); and models with joint inventory-pricing control (cf. Gong, Chao and Zheng (2014), Zhou, Tao and Chao (2014)). For continuous review models, we refer the reader to Moinzadeh and Nahmias (1988), Moinzadeh and Schmidt (1991), Bradley (2004), Song and Zipkin (2009), and the references therein.

As an exact solution seems out of reach, the operations research and management communities have instead investigated certain structural properties of the optimal policy (cf. Hua et al. (2014)), and exerted considerable effort towards constructing various heuristic policies. Veeraraghavan and Scheller-Wolf (2008) proposed the family of dual index (DI) policies, which have two base-stock levels, one for the regular source and one for the express
source, and “orders up” to bring appropriate notions of inventory position up to these levels. Scheller-Wolf, Veeraraghavan and van Houtum (2008) analyzed the closely related class of single index (SI) policies, for which the relevant notions of inventory position are different. Both families of policies seem to perform well in numerical studies. Sheopuri, Janakiraman and Seshadri (2010) considered two generalized classes of policies: one with an order-up-to structure for the express source, and one with an order-up-to structure for the regular source. Their numerical experiment showed that such policies can outperform DI policies. In the presence of production capacity costs, Boute and Van Mieghem (2014) studied dual-sourcing smoothing policies, under which the order quantities from both sources in each period are convex combinations of observed past demands. They analyzed such policies under normally distributed demand, and their numerical results showed that these policies performed better for higher capacity costs and longer lead time differences (between the two sources).

A simple and natural policy that is implemented in practice, which will be the subject of our own investigations, is the so-called Tailored Base-Surge (TBS) policy. It was first proposed and analyzed in Allon and Van Mieghem (2010), where we note that closely related standing order policies had been studied previously (cf. Rosenshine and Obee (1976), Janssen and De Kok (1999)). Under such a TBS policy, a constant order is placed at the regular source in each period to meet a base level of demand, while the orders placed at the express source follow an order-up-to rule to manage demand surges. We refer to Mini-Case 6 in Van Mieghem (2008) for more about the motivation and background of TBS policies. Note that dual-sourcing inventory systems in which a constant-order policy is implemented for the regular source are essentially equivalent to single-sourcing inventory systems with constant returns, which have been investigated in the literature (cf. Fleischmann and Kuik (2003), DeCroix, Song and Zipkin (2005)). Allon and Van Mieghem (2010) analyzed TBS policies in a continuous review model, and their focus was to find the best TBS policy. Numerical results in Klosterhalfen, Kiesmüller and Minner (2011), Rossi, Riijkema and van der Vorst (2012) showed that TBS policies are comparable to DI
policies, and outperform DI policies for some problem instances. 

Conjectured that this policy performs more effectively as the lead time difference between the two sources grows. Janakiraman, Seshadri and Sheopuri (2014) analyzed a periodic review model and studied the performance of TBS policy. They provided an explicit bound on the performance of TBS policies compared to the optimal one when the demand had a specific structure, and provided numerical experiments suggesting that the performance of the TBS policy improves as the lead time difference grows large.

However, to date there is no theoretical justification for the good behavior of TBS policies as the lead time difference grows large, and giving a solid theoretical foundation to this observed phenomena remains a major open question. We note that until recently, a similar state of affairs existed regarding the good performance of constant-order policies as the lead time grows large in single-source lost-sales inventory models. However, using tools from applied probability, queuing theory, and convexity, this phenomena was recently explained in Goldberg et al. (2012) and Xin and Goldberg (2014), in which it was proven that a simple constant-order policy is asymptotically optimal in this setting as the lead time of the single source grows large. The intuition here is that as the lead time grows large, so much randomness is introduced into the system between when an order is placed and when that order is received, that it is essentially impossible for any algorithm to meaningfully use the state information to make significantly better decisions. Thus a policy which ignores the state information (i.e. constant-order policy) performs nearly as well as an optimal policy. We note that the results of Xin and Goldberg (2014) further demonstrate that the optimality gap of the constant-order policy actually shrinks exponentially fast to zero as the lead time grows large, and provide explicit and effective bounds even for moderate-to-small lead times.

1.1. Our contributions

In this paper, we resolve this open question by proving that, when the lead time of the express source is held fixed, a simple TBS policy is asymptotically optimal as the lead time of the regular source
grows large. Our results provide a solid theoretical foundation for the conjectures and numerical experiments of Allon and Van Mieghem (2010) and Janakiraman, Seshadri and Sheopuri (2014). Interestingly, the simple TBS policy performs nearly optimally exactly when standard DP-based methodologies become intractable due to the aforementioned “curse of dimensionality”. Furthermore, as the “best” TBS policy can be computed by solving a convex program that does not depend on the lead time of the regular source (cf. Janakiraman, Seshadri and Sheopuri (2014)), our results lead directly to very efficient algorithms (with complexity independent of the lead time of the regular source) with asymptotically optimal performance guarantees. Perhaps most importantly, since many companies are already implementing such TBS policies (cf. Allon and Van Mieghem (2010)), our results provide strong theoretical support for the widespread use of TBS policies in practice. Our main proof technique combines a steady-state approach, novel convexity and lower-bounding arguments, a certain interchange of limits result, and ideas from the theory of random walks and queues, significantly extending the methodology and applicability of a novel framework for analyzing inventory models with large lead times recently introduced in Goldberg et al. (2012) and Xin and Goldberg (2014) in the context of lost-sales models with positive lead times.

1.2. Outline of paper

The rest of the paper is organized as follows. We formally define the dual-sourcing problem in Section 2 and describe the TBS policy in Section 2.1. We state our main result in Section 2.2 and prove our main result in Section 3. We summarize our main contributions and propose directions for future research in Section 4. We also include a technical appendix in Section 5.

2. Model description, problem statement and assumptions

In this section, we formally define our dual-sourcing inventory problem, closely following the definitions given in Sheopuri, Janakiraman and Seshadri (2010). Let \( \{D_t\}_{t \geq 1}, \{D'_t\}_{t \geq 1} \) be a mutually independent sequence of nonnegative independent and identically distributed (i.i.d.) demand realizations, distributed as the random variable (r.v.) \( D \). Let \( T \) be the time horizon, \( L \geq 1 \) be the deterministic lead time of the regular source (R), and \( L_0 \geq 0 \) the deterministic lead time of the
express source (E), where \( L \geq L_0 + 1 \). Let \( c_R, c_E \) be the unit purchase costs of the regular and express sources, and \( h, b \) be the unit holding and backorder costs respectively, with \( c \triangleq c_E - c_R > 0 \). In addition, let \( I_t \) denote the on-hand inventory at the start of period \( t \) (before any orders or demands are received), and \( q_t^R(q_t^E) \) denote the order placed from R(E) at the beginning of period \( t \). Note that due to the leadtimes, the order received from R(E) in period \( t \) is \( q_t^R(q_t^E - L_0) \). As we will be primarily interested in the corresponding long-run-average problem, we without loss of generality (w.l.o.g.) suppose that the initial conditions are such that the initial inventory is 0, and no initial orders have been placed from either R or E. As a notational convenience, we define \( q_k^R = 0, k = -(L-1), \ldots, 0 \); and \( q_k^E = 0, k = -(L_0-1), \ldots, 0 \). For \( t = 1, \ldots, T \), the events in period \( t \) are ordered as follows.

- Ordering decisions from R and E are made (i.e. \( q_t^E, q_t^R \) are chosen);
- New inventory \( q_t^R + q_t^E - L_0 - D_t \) is delivered and added to the on-hand inventory;
- The demand \( D_t \) is realized, costs for period \( t \) are incurred, and the inventory is updated.

Note that the on-hand inventory is updated according to \( I_{t+1} = I_t + q_t^R + q_t^E - L_0 - D_t \), and may be negative since backorder is allowed. We require that the new orders \( q_t^R \) and \( q_t^E \) are non-negative measurable (and thus deterministic) functions of the realized demands, inventory levels, and ordering quantities in periods \( 1, \ldots, t - 1 \), as well as the problem primitives \( D, L, L_0, c_R, c_E, h, b \) and the current time \( t \). We call the corresponding family of policies admissible, and denote this family by \( \Pi \). We note that any policy \( \pi \in \Pi \) can in principle be implemented on such a problem of any time horizon (even infinite). Let \( G(y) \) be the sum of the holding and backorder costs when the inventory level equals \( y \) in the end of a time period, i.e. \( G(y) \triangleq hy^+ + by^- \), where \( x^+ \triangleq \max(x, 0) \), \( x^- \triangleq \max(-x, 0) \). Here we note that \( G \) is convex and Lipschitz, and for \( x, y \in \mathbb{R} \),

\[
|G(x) - G(y)| \leq \max(b, h)|x - y|, \text{ and } |G(x)| \geq \min(b, h)|x|.
\]

Let \( C_t \) be the sum of the ordering, holding and backorder costs incurred in time period \( t \), i.e. \( C_t \triangleq c_R q_t^R + c_E q_t^E + G(I_t + q_t^R - L + q_t^E - L_0 - D_t) \). To denote the dependence of the cost on the policy \( \pi \), we use the notation \( C_t^\pi \). Let \( C(\pi) \) denote the long-run average cost incurred by a policy \( \pi \), i.e.
The value of the corresponding long-run average cost dual-sourcing inventory optimization problem is denoted by $\text{OPT}(L) \triangleq \inf_{\pi \in \Pi} C(\pi)$.

Before proceeding, it will be useful to review a certain well-known reduction between the setting in which $L_0 > 0$ and the setting in which $L_0 = 0$ (cf. Sheopuri, Janakiraman and Seshadri, 2010), where we note that similar reductions are known to hold for many classical inventory problems with backlogging (cf. Karlin and Scarf (1958), Scarf (1960)). Let us define the so-called expedited inventory position at time $t \geq 1$ as $\hat{I}_t \triangleq I_t + \sum_{k=t-L_0}^{t-1} q^E_k + \sum_{k=t-L}^{t-L_0} q^R_k$, which corresponds to the net inventory at the start of period $t$ plus all orders to be received in periods $t, \ldots, t + L_0$ (which were placed before period $t$), and the truncated regular pipeline at time $t$ as the $(L - L_0 - 1)$-dimensional vector $\mathcal{R}^t \triangleq (q^R_{t-L+L_0+1}, \ldots, q^R_{t-L_0})$, with $\mathcal{R}^t_k = q^R_{t-L+L_0+k}, k = 1, \ldots, L - L_0 - 1$. Let $\hat{\Pi}$ denote those policies belonging to $\pi$ with the additional restriction that the new orders $q^R_t, q^E_t$ are measurable functions of only $\hat{I}_t, \mathcal{R}^t$, as well as the problem primitives $D, L, L_0, c^R, c^E, h, b$ and current time $t$.

**Lemma 1** (Sheopuri, Janakiraman and Seshadri, 2010, Lemma 2.1). $\inf_{\pi \in \Pi} C(\pi) = \inf_{\pi \in \hat{\Pi}} C(\pi)$, i.e. one may w.l.o.g. restrict oneself to policies belonging to $\hat{\Pi}$.

For the remainder of the paper, we thus consider the relevant optimization only over policies belonging to $\hat{\Pi}$, i.e.

$$\inf_{\pi \in \hat{\Pi}} C(\pi).$$

Recall that a stationary Markov policy is one that places orders only based on the current state information (i.e. $\hat{I}_t$ and $\mathcal{R}^t$), but independent of the current time period $t$ and process history. It is generally well-known that for many inventory problems of interest, the relevant long-run-average optimization problems admit optimal stationary Markov policies, where such results typically follow from the general theory of infinite-horizon Markov decision processes (cf. Sennott (1989), Schäl (1993)). Explicit sufficient conditions for a rich family of inventory models to admit such an optimal policy were given in Huh, Janakiraman and Nagarajan (2011), where these conditions were verified to hold for many models of interest (e.g. lost-sales inventory models...
with positive lead times). Furthermore, it was commented in Sheopuri, Janakiraman and Seshadri (2010) that although Huh, Janakiraman and Nagarajan (2011) does not explicitly verify that their conditions hold for the dual-sourcing problem, the relevant results still hold, and a proof was sketched under some additional technical conditions. A proof was also sketched in Hua et al. (2014) under the assumption that demand is bounded. For simplicity and clarity of exposition, in the remainder of this paper we simply assume the existence of such an optimal stationary Markov policy for Problem 2. Furthermore, we also assume that of these optimal stationary policies, there exists at least one whose corresponding induced Markov chain converges in distribution and in expectation to a unique stationary measure. Again, the existence of such optimal policies is to be expected from the basic theory of Markov chains and Markov decision processes, and we refer the reader to Sennott (1989), Schäl (1993), Puterman (1994), Asmussen (2003), Meyne and Tweedie (2009), Prieto-Rumeau and Hernández-Lerma (2012), as well as the excellent recent survey of Arapostathis et al. (1993), for further details. We also note that the related work of Xin and Goldberg (2014) made a similar assumption in the context of lost-sales models. Although our precise assumptions could in principle be relaxed, e.g. to only requiring that for each $\epsilon > 0$ the stated assumptions hold for some (possibly randomized) policy which is $\epsilon$-close to optimal (as opposed to an exactly optimal deterministic policy), we do not pursue such a generalization here for the sake of brevity and clarity of exposition.

**Assumption 1.** Problem 2 has an optimal stationary Markov policy for all $L$, whose corresponding induced Markov chain converges in distribution and in expectation to a stationary measure when the initial inventory is 0, and no initial orders have been placed from either $R$ or $E$. Also, we require that $D$ is non-negative and integrable, satisfying $\mathbb{E}[D] < \infty$, and non-degenerate (i.e. not w.p.1 equal to its mean).

Let $\pi^{L,*}$ be such an optimal policy, $\left(\hat{I}^{L,*}, R^{L,*}\right)$ be a vector distributed as the stationary measure of the corresponding Markov Chain (with all r.v.s constructed on a common probability space with the appropriate joint distribution, independent of $\{D_t\}_{t \geq 1}, \{D_t'\}_{t \geq 1}$), and $D_{[t]} \triangleq (D_1, \ldots, D_t)$, where
$D_{[0]}$ denotes the empty set $\emptyset$. Let $q_{t}^{L,*} = q_t^{L,*}(\hat{I}^{L,*}, \mathcal{R}^{L,*}, D_{[t-1]})$ denote the quantity ordered from E by $\pi^*$ in period $t$ if the initial expedited inventory position equals $\hat{I}^{L,*}$, the initial truncated regular pipeline equals $\mathcal{R}^{L,*}$, and the first $t-1$ demands equal $D_{[t-1]}$. As we will be interested primarily in the setting that $L \to \infty$ with $L_0, c_R, c_E, b, h$ held fixed, we will generally suppress notationally dependence on these parameters, only making the dependence on $L$ explicit. For two r.v.s $X, Y$, let $X \sim Y$ denote equivalence in distribution between $X$ and $Y$. It follows from stationarity that

$$
\text{OPT}(L) = c_R \mathbb{E}[\mathcal{R}_1^{L,*}] + c_E \mathbb{E}[q_1^{L,*}] + \mathbb{E}
\left[
G
\left(\hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i \right)
\right];
$$

(3)

$$
\mathbb{E}[\mathcal{R}_1^{L,*}] + \mathbb{E}[q_1^{L,*}] = \mathbb{E}[D_1].
$$

(4)

Combining (3) and (4), and w.l.o.g. assuming $c_R = 0$ (cf. Sheopuri, Janakiraman and Seshadri (2010)), we have

$$
\text{OPT}(L) = c \left( \mathbb{E}[D] - \mathbb{E}[\mathcal{R}_1^{L,*}] \right) + \mathbb{E}
\left[
G
\left(\hat{I}^{L,*} + q_1^{L,*} - \sum_{i=1}^{L_0+1} D_i \right)
\right].
$$

(5)

### 2.1. TBS policy

In this section, we formally introduce the family of TBS policies, and characterize the “best” TBS policy. As a notational convenience, let us define all empty sums to equal zero, empty products to equal one, and $I(A)$ denote the indicator of the event $A$. A TBS policy $\pi_{r,S}$ with parameters $(r, S)$ is defined (cf. Janakiraman, Seshadri and Sheopuri (2014)) as the policy that places a constant order $r$ from $R$ in every period, and follows an order-up-to rule from $E$ which in each period raises the expedited inventory position to $S$ (if it is below $S$), and otherwise orders nothing. More formally, under this policy $q_t^R = r$, and $q_t^E = \max(0, S - \hat{I}_t)$, for all $t$.

Let $I^\infty(r) \overset{\Delta}{=} \sup_{t \geq 0} \left( jr - \sum_{i=1}^{j} D_i \right)$. In that case, it follows from the results of Janakiraman, Seshadri and Sheopuri (2014) that

$$
C(\pi_{r,S}) = c(\mathbb{E}[D] - r) + \mathbb{E}
\left[
G
\left(I^\infty(r) + S - \sum_{i=1}^{L_0+1} D_i \right)
\right].
$$

(6)

Note that for each $r$, the minimization problem $\inf_{S \in \mathbb{R}} C(\pi_{r,S})$ is equivalent to a standard one-period newsvendor problem. Furthermore, defining $F^\infty(r) \overset{\Delta}{=} \min_{S \in \mathbb{R}} C(\pi_{r,S})$, it is proven in
Janakiraman, Seshadri and Sheopuri (2014) that $F^\infty(r)$ is convex in $r$ on $(-\infty, E[D])$. Combining the above with standard results for single-server queues (cf. Asmussen (2003)) and (1), we conclude that there exists at least one pair $(r^*, S^*)$ such that $r^* \in \arg\min_{0 \leq r \leq E[D]} F^\infty(r)$ and $S^* \in \arg\min_{S \in \mathbb{R}} C(\pi_r, s)$; that this pair defines the TBS policy with least long-run-average cost; and that this pair can be computed efficiently by solving a convex program which is independent of the larger lead time $L$.

**2.2. Main result**

Our main result shows that the best TBS policy is asymptotically optimal as $L \to \infty$.

**Theorem 1.** Under Assumption 1, $\lim_{L \to \infty} \text{OPT}(L) = C(\pi_{r^*}, S^*)$.

**3. Proof of Theorem 1**

3.1. Lower bound for the optimal cost

In this section, we prove a lower bound for $\text{OPT}(L)$ by extending the steady-state/convexity approach of Xin and Goldberg (2014) to the dual-sourcing setting. We note that here our lower bound will involve a non-trivial optimization over measurable functions, in contrast to the bounds used in Xin and Goldberg (2014) which were of a static nature. From stationarity, for each $k = 1, \ldots, L - L_0$,

$$
\hat{I}^L_{r^*} + \sum_{i=1}^{k-1} (q_{L,r^*}^i + \mathcal{R}_{L,r^*}^i - D_i) + q_{k}^L - \sum_{i=k}^{k+L_0} D_i \sim \hat{I}^L_{r^*} + q_{L,r^*}^0 - \sum_{i=1}^{L_0+1} D_i;
$$

and for each $k = 1, \ldots, L - L_0 - 1$,

$$
\mathbb{E}[\mathcal{R}_k^L] = \mathbb{E}[\mathcal{R}_1^L] \overset{\Delta}{=} r_L.
$$

Combining the above with (5) implies that for any $\alpha \in (0, 1)$,

$$
\text{OPT}(L) \geq c(\mathbb{E}[D] - r_L) + \frac{1 - \alpha}{1 - \alpha^L} \sum_{k=1}^{L} \alpha^{k-1} \mathbb{E} \left[ G \left( \hat{I}^L_{r^*} + q_{L,r^*}^0 - \sum_{i=1}^{L_0+1} D_i \right) \right] \\
\geq c(\mathbb{E}[D] - r_L) + (1 - \alpha) \sum_{k=1}^{L - L_0} \alpha^{k-1} \mathbb{E} \left[ G \left( \hat{I}^L_{r^*} + \sum_{i=1}^{k-1} (q_{L,r^*}^i + \mathcal{R}_{L,r^*}^i - D_i) + q_{L,r^*}^k - \sum_{i=k}^{k+L_0} D_i \right) \right].
$$

Here we have introduced the discount factor $\alpha$ to implement the so-called “vanishing discount factor” approach to analyzing infinite-horizon Markov decision processes (MDP) (cf.
Huh, Janakiraman and Nagarajan (2011)), which will allow for a simpler analysis when we pass to the limit as $L \to \infty$. Indeed, this discount factor will help us to analyze the lower bound which arises when we apply Jensen’s inequality, as this lower bound will itself involve the solution to a non-trivial multi-stage dynamic optimization problem. We note that the lower bound which arose when related techniques were applied to single-sourcing systems with lost sales in Xin and Goldberg (2014) only involved a static optimization problem, and thus no such discount factor was introduced.

It then follows from the independence structure of the relevant r.v.s, and the measurability properties of $q^{L,\ast}_k$, that for each $k = 1, \ldots, L - L_0$,

$$
\mathbb{E} \left[ \hat{I}^{L,\ast} + \sum_{i=1}^{k-1} (q^{L,\ast}_i + R^{L,\ast}_i - D_i) + q^{L,\ast}_k - \sum_{i=k}^{k+L_0} D_i \bigg| D_{[k+L_0]} \right]
$$

equals

$$
\mathbb{E} [\hat{I}^{L,\ast}] + \sum_{i=1}^{k-1} (\mathbb{E}[q^{L,\ast}_i | D_{[i-1]}] + r_L - D_i) + \mathbb{E}[q^{L,\ast}_k | D_{[k-1]}] - \sum_{i=k}^{k+L_0} D_i.
$$

Further combining with the convexity of $G$ and Jensen’s inequality for conditional expectations (which applies due to Assumption 1), we obtain the following result.

**Proposition 1.** For any $\alpha \in (0, 1)$ and $L \geq L_0 + 1$, $OPT(L) - c (\mathbb{E}[D] - r_L)$ is at least

$$
(1 - \alpha) \sum_{k=1}^{L-L_0} \alpha^{k-1} \mathbb{E} \left[ G \left( \mathbb{E}[\hat{I}^{L,\ast}] - (L_0 + 1)r_L + \sum_{i=1}^{k-1} (\mathbb{E}[q^{L,\ast}_i | D_{[i-1]}] - (D_i - r_L)) \right) 
+ \mathbb{E}[q^{L,\ast}_k | D_{[k-1]}] - \sum_{i=k}^{k+L_0} (D_i - r_L) \right].
$$

(7)

Note that (7) is the discounted cost incurred (during periods $L_0 + 1, \ldots, L$) by the policy ordering $\mathbb{E}[q^{L,\ast}_i | D_{[i-1]}]$ in period $i$, of a single-sourcing $L$-period backlog inventory problem with unit holding cost $h$, backorder cost $b$, zero ordering cost, discount factor $\alpha$, i.i.d. demand distributed as $D - r_L$ (which we note can be positive or negative), lead time $L_0$, and initial inventory position (initial net inventory plus all entries of the initial pipeline vector) $\mathbb{E}[\hat{I}^{L,\ast}] - (L_0 + 1)r_L$ (cf. Karlin and Scarf (1958)), normalized by $(1 - \alpha)$. Such models, and their optimal policies, have been studied in-depth in the literature (cf. Karlin and Scarf (1958), Zipkin (2000), Fleischmann and Kuik (2003)), and are well-understood (especially for the case of non-negative demand, cf. Zipkin (2000)). Let $\hat{I}$
denote the family of all feasible non-anticipative policies for the aforementioned inventory problem (as it is typically defined, cf. Zipkin (2000)), i.e. those policies for which new orders are non-negative measurable functions of the realized demands, inventory levels, and ordering quantities in periods $1, \ldots, t-1$, as well as the current time $t$. For $\pi \in \hat{\Pi}$, initial inventory position $x \in \mathbb{R}$, $r \in \mathbb{R}$, and $i \geq 1$, let $C_i^\pi(r, x)$ denote the cost incurred by policy $\pi$ in the aforementioned inventory problem in period $i + L_0$, if the demand in each period is i.i.d. distributed as $D - r$ (with the leadtime $L_0$ and costs $b, h$ as above). For $x \in \mathbb{R}, r \in \mathbb{R}, \alpha \in (0,1), n \geq 1$, let us define

$$V_n^\alpha(r, x) \triangleq \inf_{\pi \in \hat{\Pi}} \mathbb{E}\left[\sum_{i=1}^{n} \alpha^{i-1} C_i^\pi(r, x)\right]; \quad (8)$$

and

$$V_\infty^\alpha(r, x) \triangleq \inf_{\pi \in \hat{\Pi}} \mathbb{E}\left[\sum_{i=1}^{\infty} \alpha^{i-1} C_i^\pi(r, x)\right]. \quad (9)$$

As a notational convenience, we define $V_0^\alpha(r, x) = 0$, $V_n^\alpha(r, -\infty) \triangleq \inf_{x \in \mathbb{R}} V_n^\alpha(r, x), V_\infty^\alpha(r, -\infty) \triangleq \inf_{x \in \mathbb{R}} V_\infty^\alpha(r, x)$. Then combining the above, we derive the following lower bound for $OPT(L)$.

**Lemma 2.** Under Assumption $A$ for all $\alpha \in (0,1)$ and $L \geq L_0 + 1$,

$$OPT(L) \geq c(\mathbb{E}[D] - r_L) + (1 - \alpha)V_{\infty}^{L-L_0}(r_L, -\infty). \quad (10)$$

The remainder of the proof involves demonstrating a certain interchange-of-limit results for the right-hand-side (r.h.s.) of (10). We now briefly describe the associated logic informally, and formalize all arguments in the next section. Let $r_\infty \triangleq \limsup_{L \to \infty} r_L$. Examining the r.h.s. of (10) along a subsequence on which $r_L$ converges to $r_\infty$ and taking limits, one shows that for any fixed $\alpha \in (0,1)$, $\lim_{L \to \infty} OPT(L) \geq c(\mathbb{E}[D] - r_\infty) + (1 - \alpha)V_{\infty}^{\infty}(r_\infty, -\infty)$. One then shows that the infinite-horizon problem associated with $V_{\infty}^{\infty}(r_\infty, -\infty)$ has an optimal policy which is stationary, Markov, and of order-up-to type, say to level $S_{\infty}^{\infty}(r_\infty)$. It follows that for any fixed $\alpha \in (0,1)$, $V_{\infty}^{\infty}(r_\infty, -\infty)$ is the expected infinite-horizon discounted cost incurred by the TBS policy with parameters $r_\infty, S_{\infty}^{\infty}(r_\infty)$, but possibly initialized not according to the stationary distribution of the associated inventory process, but in the state which minimizes this discounted expected infinite-horizon cost. One then appropriately bounds the difference in cost incurred under these two different
initializations uniformly in $\alpha$. Finally, letting $\alpha \uparrow 1$ will demonstrate that as $L \to \infty$, there exist TBS policies performing arbitrarily close to optimal. This will imply asymptotic optimality of the best TBS policy (with parameters $r^*, S^*$), completing the proof of our main result Theorem 1.

3.2. Interchange of limits and proof of Theorem 1

We now complete the proof of Theorem 1 by formalizing the interchange-of-limits argument sketched at the end of Section 3.1. Such interchange arguments are standard in the literature on Markov decision processes and infinite-horizon inventory control problems (cf. Iglehart (1963), Sennott (1989), Schäl (1993), Fleischmann and Kuik (2003), Feinberg (2011), Huh, Janakiraman and Nagarajan (2011)). We note that the somewhat non-standard aspect of the interchange of limits which we must demonstrate is that the demand in each period is distributed as $D - r_L$, and thus may be negative. As such, the original arguments showing such an interchange for the analogous inventory models when demand is non-negative (cf. Iglehart (1963)) do not directly apply. The possibility of negative demand also makes the verification of the conditions of general theorems which validate such an interchange (cf. Sennott (1989), Schäl (1993)) somewhat involved, even when these theorems are customized to the inventory setting (cf. Parker and Kapuscinski (2004), Huh, Janakiraman and Nagarajan (2011)). We note that the verification of closely related interchange-of-limits results have arisen recently in the context of analyzing inventory systems with returns, which reduce to standard inventory systems where demand can be positive or negative (cf. Fleischmann and Kuik (2003)). However, those results (which verify the technical conditions of Sennott (1989)) do not seem to extend immediately to our case, and further seem to require that the demand and ordering quantities take integer values. In light of the above, and for the sake of clarity and completeness, we now provide a self-contained proof of the desired interchange, which (combined with Lemma 2) will complete the proof of our main result Theorem 1.

We begin by stating some well-known properties of $V^\alpha_n(r, x)$ and $V^\infty_\alpha(r, x)$, which follow from the results of Janakiraman, Seshadri and Sheopuri (2014), Karlin and Scarf (1958) and Scarf (1960). We note that although in some cases the proofs there are only explicitly given for the case of
non-negative demand, as noted in [Heyman and Sobel (1984) and Fleischmann and Kuik (2003)],
the arguments carry over to the general case (in which demand may be negative) with only trivial
modification.

**Lemma 3** (Janakiraman, Seshadri and Sheopuri (2014), Scarf (1960)). For all \( \alpha \in (0,1), r, x \in \mathbb{R}, \) and \( n \geq 2, \)
\[
V^n_{\alpha}(r, x) = \inf_{y \geq x} \left( \mathbb{E} \left[ G(y - \sum_{k=n}^{L_0+n} (D_k - r)) \right] + \alpha \mathbb{E} \left[ V^{n-1}_{\alpha}(r, y - (D_{L_0+n} - r)) \right] \right).
\]
Furthermore, \( V^n_{\alpha}(r, x) \) is: a convex (and thus also continuous) function of \( x \) on \( \mathbb{R} \) for each fixed \( n, r; \) a continuous function of \( r \) on \( \mathbb{R} \) for each fixed \( n, x; \) an increasing function of \( x \) on \( \mathbb{R} \) for each fixed \( n, r; \) and an increasing function of \( n \) on \( \mathbb{Z}^+ \) for each fixed \( x, r. \) In addition, the infinite-horizon problem stated in the r.h.s. of (9) admits an optimal stationary Markov policy.

In preparation for demonstrating the desired interchange-of-limits, we now appropriately bound
the optimal value, and set of minimizers, of \( V^n_{\alpha}(r, x), \) uniformly in \( n. \) For \( \alpha \in (0,1) \) and \( r, x \in \mathbb{R}, \) let \( \overline{S}^\alpha_{\alpha}(r) \) denote the supremum of the set of minimizers (with respect to \( x) \) of \( V^n_{\alpha}(r, x), \) where we note that a straightforward contradiction demonstrates that \( \overline{S}^\alpha_{\alpha}(r) \in (-\infty, \infty) \) for each \( \alpha, n, r; \) and it follows from Lemma 3 that \( V^n_{\alpha}(r, -\infty) = V^n_{\alpha}(r, \overline{S}^\alpha_{\alpha}(r)). \) Then we prove the following uniform bounds, and defer the proof to the technical appendix.

**Lemma 4.** For \( \alpha \in (0,1) \) and \( r, x \in \mathbb{R}, \) \( \sup_{n \geq 1} V^n_{\alpha}(r, x) < \infty. \) Also, for each \( \alpha \in (0,1), \) there exist finite-valued strictly positive functions \( (r) \overline{S}^\alpha_{\alpha}(r), \epsilon^\alpha(r), \) which are continuous (in \( r \)) on \( \mathbb{R}, \) with
the following properties. For all \( n \geq 1: [\overline{S}^\alpha_{\alpha}(r)] < \overline{S}^\alpha_{\alpha}(r); \) and for all \( y \notin [-\overline{S}^\alpha_{\alpha}(r), \overline{S}^\alpha_{\alpha}(r)], \)
\[
\mathbb{E} \left[ G(y - \sum_{k=n}^{L_0+n} (D_k - r)) \right] + \alpha \mathbb{E} \left[ V^{n-1}_{\alpha}(r, y - (D_{L_0+n} - r)) \right] \geq V^n_{\alpha}(r, \overline{S}^\alpha_{\alpha}(r)) + \epsilon^\alpha(r).
\]

From Lemma 4 we derive the following two corollaries, whose proofs we defer to the technical appendix.

**Corollary 1.** For \( \alpha \in (0,1), \)
\[
\lim_{L \to \infty} OPT(L) \geq c(\mathbb{E}[D] - r_\infty) + (1 - \alpha) \lim_{L \to \infty} V^L_{\alpha}(r_\infty, -2\overline{S}^\alpha_{\alpha}(r_\infty)).
\]
COROLLARY 2. For all \( \alpha \in (0,1) \) and \( r, x \in \mathbb{R} \), \( V_{\alpha}^\infty(r, x) = \lim_{n \to \infty} V_{\alpha}^n(r, x) \). Furthermore, for all \( \alpha \in (0,1) \) and \( r \in \mathbb{R} \), \( V_{\alpha}^\infty(r, x) \) is a finite-valued, convex, and non-decreasing function of \( x \) on \( \mathbb{R} \). Letting \( S_{\alpha}^\infty(r) \) denote the supremum of the set of minimizers (in \( x \)) of \( V_{\alpha}^\infty(r, x) \), it holds that \( |S_{\alpha}^\infty(r)| \leq S_{\alpha}(r) \), and the infinite-horizon problem stated in the r.h.s. of (7) admits an optimal stationary base-stock policy, with order-up-to level \( S_{\alpha}^\infty(r) \).

We now formally define the Markov process representing the inventory position process under such an optimal stationary base-stock policy. Let \( S_{\alpha} \triangleq S_{\alpha}^\infty(r_\infty) \), and \( \{X_\alpha^k, k \geq 1\} \) denote the following Markov process. \( X_1^\alpha \) equals \( S_{\alpha} \). For all \( k \geq 1 \), \( X_{k+1}^\alpha = \max(X_k^\alpha + r_\infty - D_k, S_{\alpha}) \). Let \( W_k \triangleq \sum_{j=1}^k (r_\infty - D_j) \), \( Z_k \triangleq \max_{i \in [0, k-1]} W_i \), \( Z_{\infty} \triangleq \sup_{i \geq 0} W_i \), \( M_k \triangleq E[Z_k] \), \( M_\infty \triangleq E[Z_{\infty}] \). It follows from the well-known analysis of the single-server queue using Lindley’s recursion (cf. Asmussen (2003)) that \( X_k^\alpha \sim S_{\alpha} + Z_k \); and \( X_{\infty}^\alpha \lim_{k \to \infty} X_k^\alpha \) is a well-defined r.v. distributed as \( S_{\alpha} + Z_\infty \). Combining these definitions with Lemma 3 and Corollaries 1 and 2 we conclude the following.

**Corollary 3.** For \( \alpha \in (0,1) \),

\[
\lim_{L \to \infty} OPT(L) \geq c(E[D] - r_\infty) + (1 - \alpha) \sum_{k=1}^\infty \alpha^{k-1} E[G(S_{\alpha} + Z_k - \sum_{i=1}^{L_0+1} (D_i' - r_\infty))] \cdot
\]

We now briefly review some useful (and generally well-known) properties of \( Z_k \), which we will use to complete the proof of our main results, and defer the proof to the technical appendix.

**Lemma 5.** If \( r_\infty < E[D] \), then \( M_\infty < \infty \). Also, there exists \( \theta^* > 0 \) such that \( \gamma_\alpha \triangleq E[\exp(\theta^*(r_\infty - D))] \in (0,1) \), and \( M_\infty - M_0 \leq \theta^*(1 - \gamma_\alpha)^{-1} \gamma_\alpha \) for all \( n \geq 2 \). Alternatively, if \( r_\infty \geq E[D] \), then \( Z_\infty \) is almost surely infinite and \( M_\infty = \infty \). In both cases, \( \{M_k, k \geq 1\} \) is strictly increasing, \( M_\infty = \lim_{k \to \infty} M_k \), and for all \( i \geq j \geq 1 \), \( M_i - M_j = \sum_{k=j}^{i-1} k^{-1} E[\max(0, W_k)] \). Also, if \( r_\infty = E[D] \), there exists a strictly positive finite constant \( C^* \) (depending only on \( D \)) such that for all \( i \geq j+1 \geq 3 \), \( M_i - M_j \geq C^*(i^\frac{1}{2} - j^\frac{1}{2}) \).

We now combine Corollary 3 with Lemma 5 to complete the proof of our main result.

**Proof of Theorem 7** First, we prove that \( r_\infty < E[D] \), which we will need to demonstrate the desired uniform convergence in \( \alpha \). We note that this will require somewhat subtle arguing, since
we must show that otherwise (i.e. if \( r_\infty = \mathbb{E}[D] \)) the cost associated with the relevant inventory process grows too quickly (as \( \alpha \to 1 \)) even when initialized to \( S_\alpha \), which may be converging to \(-\infty\) as \( \alpha \to 1 \). Suppose for contradiction that \( r_\infty = \mathbb{E}[D] \). In this case, it follows from Corollary 3 (11), and Jensen’s inequality that for all \( \alpha \in (0,1) \),

\[
\lim_{L \to \infty} OPT(L) \geq \min(b, h) \inf_{S \in \mathbb{R}} \sum_{k=1}^{\infty} (1 - \alpha) \alpha^{k-1} |S + M_k|.
\]

Let \( G_\alpha \) denote a geometrically distributed r.v. with success probability \( 1 - \alpha \), independent of \( \{Z_k, k \geq 1\} \), and \( m_\alpha \overset{\Delta}{=} \lceil -\frac{1}{\log_2(\alpha)} \rceil \) denote a median of \( G_\alpha \). Note that the memoryless property implies \( \mathbb{P}(G_\alpha \geq 2m_\alpha) = \frac{1}{2} \), and that we may interpret the r.h.s. of (11) as an appropriate single-stage newsvendor problem (with ordering level \( S \) and demand distributed as \( M_{G_\alpha} \)). We conclude from Lemma 5 well-known results for the newsvendor problem (cf. Zipkin (2000)), and the memoryless property that for all sufficiently large \( \alpha \in (0,1) \), \( \lim_{L \to \infty} OPT(L) \) is at least

\[
\min(b, h) \mathbb{E}[|M_{2m_\alpha} - M_{G_\alpha}|] \geq \frac{1}{4} \min(b, h) (M_{2m_\alpha} - M_{m_\alpha}) \geq \frac{1}{4} \min(b, h) C^* \left( (2m_\alpha) \frac{1}{2} - m_\alpha \frac{1}{2} \right). \]

As it is easily verified that \( \lim_{\alpha \uparrow 1} \left( (2m_\alpha) \frac{1}{2} - m_\alpha \frac{1}{2} \right) = \infty \), we conclude that if \( r_\infty = \mathbb{E}[D] \), then \( \lim_{L \to \infty} OPT(L) = \infty \). However, in this case a contradiction is easily reached by considering the TBS policy \( \pi_{0,0} \), which incurs long-run average cost \( C(\pi_{0,0}) < \infty \) for all \( L \geq L_0 + 1 \), completing the proof that \( r_\infty < \mathbb{E}[D] \).

In that case, it follows from (6) that for all \( \alpha \in (0,1) \),

\[
C(\pi_{r_\infty, S_{r_\infty} + (L_0 + 1)r_\infty}) = c(\mathbb{E}[D] - r_\infty) + \mathbb{E} \left[ G \left( S_\alpha + Z_\infty + (L_0 + 1)r_\infty - \sum_{i=1}^{L_0+1} D_i \right) \right] \]

\[
= c(\mathbb{E}[D] - r_\infty) + (1 - \alpha) \sum_{k=1}^{\infty} \alpha^{k-1} \mathbb{E} \left[ G \left( S_\alpha + Z_\infty - \sum_{i=1}^{L_0+1} (D_i' - r_\infty) \right) \right].
\]

Combining with Corollary 3, Lemma 5 and (11), we conclude that for all \( \alpha \in (0,1) \), \( C(\pi^{*\alpha}, S^\alpha) - \lim_{L \to \infty} OPT(L) \) is at most

\[
(1 - \alpha) \sum_{k=1}^{\infty} \alpha^{k-1} \left( \mathbb{E} \left[ G \left( S_\alpha + Z_\infty - \sum_{i=1}^{L_0+1} (D_i' - r_\infty) \right) \right] - \mathbb{E} \left[ G \left( S_\alpha + Z_k - \sum_{i=1}^{L_0+1} (D_i' - r_\infty) \right) \right] \right) \]

\[
\leq \max(b, h)(1 - \alpha) \left( M_\infty + \sum_{k=1}^{\infty} \alpha^{k-1} (\theta^* (1 - \gamma^*_k))^{-1} \gamma^*_k \right) \]

\[
= \max(b, h) \left( (1 - \alpha) M_\infty + \gamma^*_k (\theta^* (1 - \gamma^*_k))^{-2} \frac{1 - \alpha}{1 - \gamma^*_k} \right). \quad (12)
\]
Noting that \(12\) converges to 0 as \(\alpha \uparrow 1\) completes the proof. □

4. Conclusion

In this paper, we proved that when the lead time of the express source is held fixed, a simple TBS policy is asymptotically optimal for the dual-sourcing inventory problem as the lead time of the regular source grows large. Our results provide a solid theoretical foundation for several conjectures and numerical experiments appearing previously in the literature regarding the good empirical performance of such policies. Furthermore, the simple TBS policy performs nearly optimally exactly when standard DP-based methodologies become intractable due to the curse of dimensionality. In addition, since the “best” TBS policy can be computed by solving a convex program that does not depend on the lead time of the regular source, and is easy to implement, our results lead directly to very efficient algorithms with asymptotically optimal performance guarantees. Perhaps most importantly, since many companies are already implementing such TBS policies, our results provide strong theoretical support for the widespread use of TBS policies in practice.

This work leaves many interesting directions for future research. First, it would be interesting to investigate the rate of convergence to optimality of TBS policies as the lead time grows large, especially in light of their use in practical settings. Such an analysis would seem to involve estimates for the rate of convergence of finite horizon inventory optimization problems to their infinite horizon counterparts, which has been previously investigated for related systems (cf. Hordijk and Tijms (1974, 1975)). We do note that a detailed analysis of the proofs of our main results in principle yields a computable bound for this rate. For example, one can minimize the r.h.s. of \(10\) over \(r_L \in [0, \mathbb{E}[D]]\), and maximize over \(\alpha \in (0,1)\), to yield a (relatively) easy-to-compute bound for each \(L \geq L_0 + 1\) (which can be further improved by more carefully optimizing our approach and bounds for this purpose). However, a more precise theoretical analysis of the performance of TBS policies for small-to-moderate lead times, analogous to the exponential rate of convergence to optimality of the constant-order policy for lost-sales inventory models identified in Xin and Goldberg (2014), seems to require fundamentally new ideas.
Second, and related to the aforementioned discussion as regards the rate of convergence to optimality of TBS policies, it would be interesting to identify other more sophisticated algorithms which perform better for small-to-moderate lead times, yet remain efficient to implement. Indeed, it remains an interesting open question to better understand the trade-off between algorithmic run-time and achievable performance guarantees in this context, i.e. how complex an algorithm is required to “exploit” the weak correlations which persist even as the lead time grows large. In the context of dual-sourcing, potential algorithms here include: the so-called dual-sourcing smoothing policies recently studied in Boute and Van Mieghem (2014); affine policies more generally (cf. Ben-Tal et al. (2005), Bertsimas, Iancu and Parrilo (2010)), of which dual-sourcing smoothing policies are a special case; the single index and dual index policies discussed earlier; or the dual-balancing policies analyzed in Levi, Janakiraman and Nagarajan (2008). On a related note, it would be quite interesting to analyze “hybrid” algorithms, which could e.g. solve a large dynamic program when the lead time is small, and gradually transition to using simpler heuristics as the lead time grows large; or combine different heuristics depending on the specific problem parameters.

On a final note, combined with the results of Goldberg et al. (2012) and Xin and Goldberg (2014), our methodology lays the foundations for a completely new approach to analyzing inventory models with large lead times. So far, this approach has been successful in yielding key insights and efficient algorithms for two settings previously believed intractable: lost sales models with large lead times, and dual-sourcing models with large lead time gap. We believe that our techniques have the potential to make similar progress on many other difficult supply chain optimization problems of practical relevance in which there is a lag between when policy decisions are made and when those decisions are implemented. This includes both more realistic variants of the lost-sales and dual-sourcing models considered so far (e.g. models with distributional dependencies, parameter uncertainty, complex network structure, and more accurate modeling of costs), as well as fundamentally different models (e.g. inventory systems with remanufacturing when the manufactured and remanufactured lead times differ, cf. Zhou, Tao and Chao (2011); multi-echelon systems with
lost sales and positive lead times, cf. [Huh and Janakiraman (2010)]; or models with perishable
goods). In closing, we note that our approach can more generally be viewed as a methodology
to formalize the notion that when there is a high level of uncertainty and randomness in one’s
supply chain, even simple policies perform nearly as well as very sophisticated policies, since no
algorithm can “beat the noise”. Exploring this concept from a broader perspective may be fruitful
in yielding novel algorithms and insights for a multitude of problems in operations management
and operations research.

Acknowledgments
The authors thank Shabbir Ahmed, Xin Chen, Jim Dai, Alan Erera, Ganesh Janakiraman, Retsef
Levi, Alan Scheller-Wolf, Alex Shapiro,Cong Shi, Mark Squillante, and Qiong Wang for several
stimulating discussions. We also would like to give a special thanks to Ganesh Janakiraman for
pointing out several references to the literature on inventory systems with returns.

References
Allon, G., J. A. Van Mieghem. 2010. Global Dual Sourcing: Tailored Base-Surge Allocation to Near- and
Offshore Production. Management Science. 56(1) 110-124.

Angelus, A., Ö. Özer, 2015. Knowledge you can act on: optimal policies for assembly systems with expediting
and advance demand information. working paper.

Arapostathis, A., V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, S. I. Marcus. 1993. Discrete time
controlled Markov processes with average cost criterion: a survey. SIAM Journal on Control and Opti-
mization. 31 282-344.

Asmussen, S. 2003. Applied probability and queues. Second ed., Springer.

Barankin, E. W. 1961. A delivery-lag inventory model with an emergency provision. Naval Research Logistics
Quarterly. 8(3) 285-311.

Ben-Tal, A., B. Golany, A. Nemirovski, J.P. Vial. 2005. Retailer-supplier flexible commitments contracts: a
robust optimization approach. Manufacturing & Service Operations Management. 7 248–271.

Ben-Tal, A., A. Goryashko, A. Guslitzer, A. Nemirovski. 2004. Adjustable robust solutions of uncertain linear
programs. Mathematical Programming. 99 351-376.
Bertsimas, D., D. A. Iancu, P.A. Parrilo. 2010. Optimality of affie policies in multistage robust optimization. 

*Mathematics of Operations Research.** 35**(2) 363-394.

Bradley, J. R. 2004. A Brownian approximation of a production-inventory system with a manufacturer that subcontracts. *Operations Research.* 52**(5) 765-784.

Boute, R.N., J.A. Van Mieghem. 2014. Global dual sourcing and order smoothing: The impact of capacity and leadtimes. *Management Science.* forthcoming.

Chen, Y., W. Xue, J. Yang. 2013. Optimal inventory policy in the presence of a long-term supplier and a spot market. *Operations Research.* 61**(1) 88-97.

Daniel, K. H. 1963. A delivery-lag inventory model with emergency order. H. Scarf, D. Gilford, M. Shelly, eds. *Multistage Inventory Models and Techniques,* chapter 2. Stanford University Press, Stanford, CA.

DeCroix, G., J.S. Song, P. Zipkin. 2005. A series system with returns: Stationary analysis. *Operations Research.* 53 350-362.

Erdos, P., M. Kac. 1946. On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* 52 292-302.

Feinberg, E.A. 2011. Total expected discounted reward MDPs: existence of optimal policies. *Wiley Encyclopedia of Operations Research and Management Science.*

Feng, Q., S. Sethi, H. Yan, H. Zhang. 2006. Are base-stock policies optimal in inventory problems with multiple delivery modes? *Operations Research.* 54**(4) 801-807.

Fleischmann, M., R. Kuik. 2003. An optimal inventory control with independent stochastic item returns. *European Journal of Operational Research.* 151 25-37.

Folland, G. B. 1999. *Real Analysis: Modern Techniques and Their Applications.* Second Edition, Wiley.

Fox, E.J., R. Metters, J. Semple. 2006. Optimal inventory policy with two suppliers. *Operations Research.* 54**(2) 389-393.

Fukuda, Y. 1964. Optimal policies for the inventory problem with negotiable leadtime. *Management Science.* 10 690-708.
Goldberg, D.A., D. A. Katz-Rogozhnikov, Y. Lu, M. Sharma, M. S. Squillante. 2012. Asymptotic optimality of constant-order policies for lost sales inventory models with large lead times. Second round of review at Mathematics of Operations Research. [http://arxiv.org/abs/1211.4063](http://arxiv.org/abs/1211.4063).

Gong, X., X. Chao, S. Zheng. 2014. Dynamic pricing and inventory management with dual suppliers of different lead times and disruption risks. *Production and Operations Management*. forthcoming.

Heyman, D., M. Sobel. 1984. *Stochastic Models in Operations Research*, Vol. II. McGraw-Hill, New York.

Hordijk, A., H. Tijms. 1974. Convergence Results and Approximations for Optimal (s, S) Policies. *Management Science*. 20(11) 1432-1438.

Hordijk, A., H. Tijms. 1975. On a conjecture of Iglehart. *Management Science*. 21(11) 1342-1345.

Hua, Z., Y. Yu, W. Zhang, X. Xu. 2014. Structural properties of the optimal policy for dual-sourcing systems with general lead times. *IIE Transactions*. forthcoming.

Huggins, E. L., T. L. Olsen. 2010. Inventory control with generalized expediting. *Operations Research*. 58(5) 1414-1426.

Huh, W.T., G. Janakiraman. 2010. On the optimal policy structure in serial inventory systems with lost sales. *Operations Research*. 58(2) 486-491.

Huh, W.T., G. Janakiraman, M. Nagarajan. 2011. Average cost single-stage inventory models: an analysis using a vanishing discount approach. *Operations Research*. 59(1) 143-155.

Iglehart, D. L. 1963. Optimality of (s, S) policies in the infinite horizon dynamic inventory problem. *Management Science*. 9(2) 259-267.

Janakiraman, G., S. Seshadri, A. Sheopuri. 2014. Analysis of Tailored Base-Surge Policies in Dual Sourcing Inventory Systems. *Management Science*. forthcoming.

Janssen, F., T. De Kok. 1999. A two-supplier inventory model. *International Journal of Production Economics*. 59 395-403.

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.
Kelleher, K. 2003. Why FedEx is gaining ground. *Business 2.0* (October) 56-57.

Kingman, J.F.C. 1962. Some inequalities for the queue GI/G/1. *Biometrika* **49** 315-324.

Klosterhalfen, S., G. Kiesmüller. S. Minner. 2011. A comparison of the constant-order and dual-index policy for dual sourcing. *International Journal of Production Economics*. **133**(1) 302-311.

Kuhn, D., W. Wiesemann, A. Georgiou. 2011. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*. **130**(1) 177-209.

Lawson, D. G., E. L. Porteus. 2000. Multistage inventory management with expediting. *Operations Research*. **48** 878-893.

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

Meyn, S., R. Tweedie. 2009. *Markov chains and stochastic stability*. Cambridge University Press.

Minner, S. 2003. Multiple-supplier inventory models in supply chain management: A review. *International Journal of Production Economics*. **81-82** 265-279.

Moinzadeh, K., S. Nahmias. 1988. A continuous review model for an inventory system with two supply modes. *Management Science*. **34**(6) 761-773.

Moinzadeh, K., C. P. Schmidt. 1991. An (S-1, S) inventory system with emergency orders. *Operations Research*. **39**(2) 308-321.

Morton, T. E. 1969. Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM Review*. **11**(4) 572-596.

Neuts, F. 1964. An inventory model with optimal time lag. *SIAM Journal on Applied Mathematics*. **12** 179-185.

Parker, R., R. Kapuscinski. 2004. Optimal policies for a capacitated twoechelon inventory system. *Operations Research*. **52** 739-755.

Prieto-Rumeau, T., O. Hernández-Lerma. 2012. *Selected topics on continuous-time controlled Markov chains and Markov games*. Imperial College Press, London.
Puterman, M. L. 1994. *Markov Decision Processes: discrete stochastic dynamic programming*. John Wiley & Sons, New York, NY.

Rao, U., A. Scheller-Wolf, S. Tayur. 2000. Development of a rapidresponse supply chain at Caterpillar. *Operations Research*. **48**(2) 189-204.

Rosenshine, M., D. Obee. 1976. Analysis of a standing order inventory system with emergency orders. *Operations Research*. **24**(6) 1143-1155.

Rossi, R., W. A. Rijpkema, J. G. A. J. van der Vorst. 2012. The impact of dual sourcing on food supply chain networks: the case of Egyptian strawberries. In *Proc. of the 11th Wageningen International Conference on Chain and Network Management*, Wageningen.

Scarf, H. 1960. The Optimality of (s, S) Policies in the Dynamic Inventory Problem. *Mathematical Methods in the Social Sciences*. Stanford University Press. 196-202.

Schäl, M. 1993. Average optimality in dynamic programming with general state space. *Mathematics of Operations Research*. **18**(1) 163-172.

Scheller-Wolf, A., S. Veeraraghavan, G.-J. van Houtum. 2008. Effective Dual Sourcing with a Single Index Policy. Working paper.

Sennott, L.I., 1989. Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs. *Operations Research*. **37**(4) 626-633.

Sethi, S. P., H. Yan, H. Zhang. 2003. Inventory models with fixed costs, forecast updates, and two delivery modes. *Operations Research*. **51**(2) 321-328.

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.

Song, J., P. Zipkin. 2009. Inventories with Multiple Supply Sources and Networks of Queues with Overflow Bypasses. *Management Science*. **55**(3) 362-372.

Spitzer, F. 1956. A combinatorial lemma and its application to probability theory. *Transactions of the American Mathematical Society*. **82**(2) 323-339.

Tagaras, G., D. Vlachos. 2001. A periodic review inventory system with emergency replenishments. *Management Science*. **47**(3) 415-429.
Tao, Z., X. S. Zhou. 2014. Approximation balancing policies for inventory systems with remanufacturing. *Mathematics of Operations Research*. forthcoming.

Van Mieghem, J. A. 2008. *Operations Strategy: Principles and Practice*. Dynamic Ideas, Belmont, MA.

Veeraraghavan, S., A. Scheller-Wolf. 2008. Now or later: Dual index policies for capacitated dual sourcing systems. *Operations Research*. 56(4) 850-864.

Whittmore, A. S., S. C. Saunders. 1977. Optimal inventory under stochastic demand with two supply options. *SIAM Journal on Applied Mathematics*. 32 293-305.

Xin, L., D.A. Goldberg. 2014. Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. arXiv preprint arXiv:1409.1499.

Yi, J., A. Scheller-Wolf. 2003. Dual sourcing from a regular supplier and a spot market. Working paper, Carnegie Mellon University, Pittsburgh.

Zhang, V. L. 1996. Ordering policies for an inventory system with three supply modes. *Naval Research Logistics*. 43 691-708.

Zhou, S.X., Z. Tao, X. Chao. 2011. Optimal control of inventory systems with multiple types of remanufacturable products. *Manufacturing & Service Operations Management*. 13(1) 20-34.

Zhou, S.X., X. Chao. 2014. Dynamic pricing and inventory management with regular and expedited supplies. *Production and Operations Management*. 23(1) 65-80.

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

Zipkin, P. 2000. Foundations of inventory management. Vol. 2. New York: McGraw-Hill.

5. Technical Appendix

*Proof of Lemma 4* By evaluating the policy which never orders, we conclude that for all $\alpha \in (0, 1)$, $r, x \in \mathbb{R}$, $\sup_{n \geq 1} V_n^\alpha(r, x)$ is at most $\mathbb{E}\left[\sum_{i=1}^{\infty} \alpha^{i-1} G(x - \sum_{j=1}^{i} (D_j - r) - \sum_{k=i+1}^{L_0+i} (D_k - r))\right]$, which by (1) is itself bounded by

$$\max(b, h)(|x| + |r| + \mathbb{E}[D]) \sum_{i=1}^{\infty} (i + L_0) \alpha^{i-1} \leq 2(L_0 + 1) \max(b, h)(|x| + |r| + \mathbb{E}[D])(1 - \alpha)^{-2}.$$
Combining with (1) and a straightforward calculation, it follows that for any \( r, y \in \mathbb{R} \) such that 
\[
|y| \geq \overline{S}_\alpha(r) \triangleq 4(L_0 + 1)\max(b, h) \left[ |r| + \mathbb{E}[D] \right] \left( 1 - \alpha \right)^{-2}, \text{ and all } n \geq 1, \text{ one has that } \mathbb{E}[G(y - \sum_{k=n}^{L_0+n} (D_k - r))] - V_n^\alpha (r, 0) \geq \epsilon_n(r) \triangleq (L_0 + 1) \max(b, h) \mathbb{E}[D]. \text{ Combining the above completes the proof.} \]

\[
\begin{proof}[Proof of Corollary 1] \]
Let \( \{i_k, k \geq 1\} \) be a strictly increasing subsequence of positive numbers such that \( i_1 \geq L_0 + 1 \), and \( \lim_{k \to \infty} r_{i_k} = r_\infty \) (existence follows from the definition of \( \limsup \)). It follows from Lemmas 2, 3, and 4 that for all \( k \geq 1 \),
\[
\text{OPT}(i_k) \geq c(\mathbb{E}[D] - r_{i_k}) + (1 - \alpha)V_{\alpha}^{i_k - L_0}(r_{i_k} - \overline{S}_\alpha(r_{i_k})).
\]
It then follows from the monotonicity (in \( n \)) of \( V_{\alpha}^n(r, x) \) that for all \( k \geq k_0 \) and \( L \leq i_k - L_0 \),
\[
\text{OPT}(i_k) \geq c(\mathbb{E}[D] - r_{i_k}) + (1 - \alpha)V_{\alpha}^L(r_{i_k} - 2\overline{S}_\alpha(r_\infty)).
\]
Fixing \( L \geq 1 \), letting \( k \to \infty \), and applying the continuity (in \( r \)) of \( V_{\alpha}^\infty(r, x) \) completes the proof. \( \Box \)

\[
\begin{proof}[Proof of Corollary 2] \]
We first demonstrate that \( V_{\alpha}^\infty(r, x) = \lim_{n \to \infty} V_{\alpha}^n(r, x) \). The existence of the corresponding limit follows from the monotonicity (in \( n \)) guaranteed by Lemma 3. That \( V_{\alpha}^\infty(r, x) \geq \lim_{n \to \infty} V_{\alpha}^n(r, x) \) for all \( \alpha \in (0, 1) \) and \( r, x \in \mathbb{R} \) follows immediately from the definitions of the associated optimization problems. To prove the other direction, we note that for any fixed \( n \geq 1 \), it follows from the convexity ensured by Lemma 3 that there exists an optimal policy \( \pi \) for the problem stated in the r.h.s. of (8) of base-stock form, with order-up-to levels \( C_1, \ldots, C_n \) (i.e. order up to level \( C_i \) in period \( i \) if the pre-order inventory level is below \( C_i \), otherwise order nothing). Furthermore, it follows from Lemma 4 that \( \max_{i=1, \ldots, n} |C_i| \leq \overline{S}_\alpha(r) \). Now, consider the policy \( \pi' \) that orders up to level \( C_i \) in period \( i \) if the pre-order inventory level is below \( C_i \) and otherwise orders nothing in periods \( i = 1, \ldots, n \); and orders nothing in all remaining periods, irregardless of
the inventory level. Note that under policy \( \pi' \), w.p.1 the absolute value of the inventory position at the end of period \( i \) is at most \( |x| + \sum_{r} s_{\alpha}(r) + i|r| + \sum_{k=1}^{i} D_k \). It then follows from an argument nearly identical to that presented in our proof of Lemma 4 that

\[
\mathbb{E}\left[ \sum_{i=1}^{\infty} \alpha^{i-1} C_i^*(r, x) \right] - V_n^\infty(r, x) \leq \max(b, h)(s_\alpha(r) + |x| + |r| + \mathbb{E}[D]) \sum_{i=n+1}^{\infty} (i + L_0) \alpha^{i-1}. \tag{13}
\]

That \( V_n^\infty(r, x) \leq \lim_{n \to \infty} V_n^\infty(r, x) \) then follows from the fact that the r.h.s. of (13) converges to 0 as \( n \to \infty \). The remainder of the corollary follows by combining the above with Lemma 4 and applying the fact that convexity and monotonicity are preserved under limits. □

Proof of Lemma 5

The entirety of the lemma, barring the lower bound involving \( C^* \), follows by combining generally well-known results for generating functions, large deviations, single-server queues, and recurrent random walks (cf. Spitzer (1956), Kingman (1962), Folland (1999), Asmussen (2003), Xin and Goldberg (2014)), and we omit the details. We now prove the lower bound involving \( C^* \), which we note would follow from well-known weak-convergence results under additional assumptions on \( D \) (e.g. finite variance, cf. Erdos and Kac (1946)). Let \( \{A_i^+, i \geq 1\} \) denote an i.i.d. sequence of r.v.s distributed as \( \mathbb{E}[D] - D \) conditioned on the event \( \{\mathbb{E}[D] > D\} \), and \( \{A_i^-, i \geq 1\} \) denote an i.i.d. sequence of r.v.s distributed as \( D - \mathbb{E}[D] \) conditioned on the event \( \{D \geq \mathbb{E}[D]\} \).

Let \( B_k \) denote a binomially distributed r.v. with parameters \( k, p \), where \( p \) is defined as \( \mathbb{P}(\{\mathbb{E}[D] > D\}) \), independent of \( \{A_i^+, i \geq 1\} \) and \( \{A_i^-, i \geq 1\} \). Note that we may construct \( W_k \) on an appropriate probability space such that \( W_k = \sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^- \), in which case (by non-negativity) \( \mathbb{E}[\max(0, W_k)] \) is at least

\[
\mathbb{E}\left[ \sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^- \right] \left\{ B_k \geq pk + (p(1-p)k)^{\frac{1}{2}} \right\} \mathbb{P}\left( \frac{B_k - pk}{(p(1-p)k)^{\frac{1}{2}}} \geq 1 \right).
\]

Furthermore, since \( p\mathbb{E}[A_i^+] = (1-p)\mathbb{E}[A_i^-] \), it follows from non-negativity and independence that

\[
\mathbb{E}\left[ \sum_{i=1}^{B_k} A_i^+ - \sum_{i=1}^{k-B_k} A_i^- \right] \left\{ B_k \geq pk + (p(1-p)k)^{\frac{1}{2}} \right\} \geq (pk + (p(1-p)k)^{\frac{1}{2}})\mathbb{E}[A_i^+] - ((1-p)k - (p(1-p)k)^{\frac{1}{2}})\mathbb{E}[A_i^-] = (p(1-p)k)^{\frac{1}{2}}(\mathbb{E}[A_i^+] + \mathbb{E}[A_i^-]).
\]
Combining the above with the central limit theorem, we conclude that

\[
\liminf_{k \to \infty} \frac{\mathbb{E}[\max(0, W_k)]}{k^{1/2}} \geq (p(1-p))^{1/2} (\mathbb{E}[A_1^+] + \mathbb{E}[A_1^-]) \liminf_{k \to \infty} \mathbb{P}\left( \frac{B_k - pk}{(p(1-p)k)^{1/2}} \geq 1 \right) > 0.
\]

It then follows from the strict positivity of \( \mathbb{E}[\max(0, W_k)] \) for all \( k \geq 2 \) that there exists \( C' > 0 \) such that \( \mathbb{E}[\max(0, W_k)] \geq C' k^{1/2} \) for all \( k \geq 2 \). Thus for all \( i \geq j \geq 2 \),

\[
M_i - M_j = \sum_{k=j}^{i-1} k^{-1} \mathbb{E}[\max(0, W_k)] \\
\geq C' \sum_{k=j}^{i-1} k^{-1/2} \geq C' \int_j^i x^{-1/2} dx = 2C'(i^{1/2} - j^{1/2}).
\]

Combining the above completes the proof. \( \square \)