

# Projected Inventory Level Policies for Lost Sales Inventory Systems: Asymptotic Optimality in Two Regimes

Willem van Jaarsveld

School of Industrial Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands, PO BOX 513, 5600MB, w.l.v.jaarsveld@tue.nl

Joachim Arts

Luxembourg Centre for Logistics and Supply Chain Management, Department of Economics and Management, University of Luxembourg, Luxembourg City, Luxembourg, 6, rue Richard Coudenhove-Kalergi L-1359 joachim.arts@uni.lu

We consider the canonical periodic review lost sales inventory system with positive lead-times and stochastic i.i.d. demand under the average cost criterion. We introduce a new policy that places orders such that the expected inventory level at the time of arrival of an order is at a fixed level and call it the Projected Inventory Level (PIL) policy. We prove that this policy has a cost-rate superior to the equivalent system where excess demand is back-ordered instead of lost and is therefore asymptotically optimal as the cost of losing a sale approaches infinity under mild distributional assumptions. We further show that this policy dominates the constant order policy for any finite lead-time and is therefore asymptotically optimal as the lead-time approaches infinity for the case of exponentially distributed demand per period. Numerical results show this policy also performs superior relative to other policies.

*Key words:* Lost Sales, Inventory, Optimal policy, Asymptotic Optimality, Markov Decision Process

---

## 1. Introduction

The periodic review inventory system with lost sales, positive lead time and i.i.d. demand is a canonical problem in inventory theory. The decision maker is interested in the average cost-rate of this system. The optimal policy for such a system can be computed in principle by stochastic dynamic programming, but it is not practical due to the curse of dimensionality. Research has therefore focused on devising heuristic policies for the lost sales inventory system. Although many

variants of lost sales inventory systems exist, results for the canonical system are important as they serve as building blocks to design good policies for more intricate lost sales inventory systems. Bijvank and Vis (2011) review the literature on many such more intricate inventory systems with lost sales.

There are two simple heuristic policies for the canonical lost sales system that appeal to both practitioners and academics. These policies are the base-stock policy and the constant order policy.

The base-stock policy places an order each period such that the inventory position (inventory on-hand + outstanding orders) is raised to a fixed *base-stock level*. This policy is prevalent in practice due to its intuitive structure and because it is the optimal policy when excess demand is not lost but back-ordered. The most important merit of the base-stock policy is that it is asymptotically optimal as the cost of a lost sale approaches infinity under mild conditions on the demand distribution (Huh et al. 2009b). This asymptotic optimality is robust in the sense that it holds for a broad class of heuristics to compute base-stock levels (Bijvank et al. 2014).

The constant order policy orders the same amount in each period regardless of the state of the system. Although this may seem naive at first, this policy is asymptotically optimal as the lead-time approaches infinity (Goldberg et al. 2016), and can outperform the base-stock policy for long lead-times and moderate costs for a lost sale. Xin and Goldberg (2016) show that the constant order policy converges to optimality exponentially fast in the lead time.

The asymptotic optimality results of Huh et al. (2009b) and Goldberg et al. (2016) are both elegant and useful for practice. We believe that these results should also inform the design of plausible heuristic policies: With the knowledge that such asymptotic optimality results are attained by relatively simple policies, new heuristics for the lost sales system should be designed to be asymptotically optimal for long lead times *and* large lost sales penalties. Unfortunately, the constant order policy is not asymptotically optimal for large lost sales penalties and the base-stock policy is not asymptotically optimal for long lead times. This paper introduces a single parameter policy that is asymptotically optimal for large lost sales penalties under mild assumptions on the demand

distribution and also for long lead times when demand has an exponential distribution. We call this policy the *projected inventory level* (PIL) policy.

The PIL policy places orders such that the expected inventory level at the time of arrival of an order is raised to a fixed level which we call the projected inventory level. The PIL policy is intuitive for academics and practitioners alike. In fact, the base-stock policy for the canonical inventory system where excess demand is back-ordered, rather than lost, is also a projected inventory level policy: Although the usual interpretation of a base-stock policy in a system with back-orders is that it raises the inventory position to a fixed level, it is equivalent to say that it raises the expected inventory level at the time of order arrival to a fixed level. (These two policies are not equivalent in the lost sales inventory system.) We exploit this equivalence and use it to compare the canonical inventory systems with lost sales and back-orders respectively when all parameters are identical. We prove that the cost-rate for the canonical lost sales system is lower than the cost-rate for the canonical back-order system under the same projected inventory level. As a corollary to this, we find that optimal cost-rate of the canonical lost sales system is lower than the optimal cost-rate for the equivalent canonical back-order system. This means we recover a main result of Janakiraman et al. (2007) via a new and different proof. The stochastic comparison technique used by Janakiraman et al. (2007) holds for general convex per period cost functions but does not construct a policy. Our result is *constructive* in the sense that we identify a specific policy (the PIL policy) for which the costs in the lost sales system are lower than the optimal costs for the back-order system. Our construction requires that the cost per period is linear in on-hand inventory and lost sales, which is the most commonly used per period cost function.

The PIL policy also mimics the behavior of the constant order policy for long lead-times. We make this notion rigorous when demand per period has an exponential distribution. In that case, we show that the projected inventory level policy can be interpreted as a one-step policy improvement on the (bias function of) the constant order policy; we believe this to be an interesting proof technique. Under the same assumption, we show that the projected inventory level policy dominates

the constant order policy for any finite lead-time  $\tau$ . The PIL policy therefore inherits the property of the constant order policy that the gap with the optimal policy decreases exponentially with the lead-time cf. Xin and Goldberg (2016).

Note that the projected inventory level policy has a single parameter and yet uses all the information in the state vector without aggregating it into the inventory position. This is a feature shared by the Myopic policy where orders are placed to minimize the expected cost in the period that the order arrives. These policies all require a projection of the inventory level at the time of order arrival. The myopic and PIL policy can therefore both be considered as “projection” policies. Empirically projection policies perform exceptionally well (see Zipkin (2008a) and Section 7), but there is no theoretical underpinning that explains why such policies perform so well empirically. In particular, there are no known asymptotic optimality results for such policies. This paper contributes asymptotic optimality results for the projected inventory level policy in two asymptotic regimes.

Policies that are asymptotically optimal for high lost sales costs and long lead times can be constructed by using multiple parameters to make the policy behave as either a base-stock policy or a constant order policy when needed. For example, a policy may order a convex combination of the base-stock and constant order policy decision, or cap the order placed by a base-stock policy Johansen and Thorstenson (2008), Xin (2020). The projected inventory level policy is different in that the parameter can not be set such that it trivially reduces to either a base-stock or constant order policy. Asymptotic optimality proofs therefore rely on new ideas.

In summary, this paper makes the following contributions:

1. We introduce the projected inventory level policy and show that it is a natural generalization of the base-stock policy to systems where sales are lost rather than back-ordered. Furthermore, it is a single parameter policy that utilizes all state information without aggregating it into the inventory position, i.e., it is a projection policy.
2. We provide the first tractable policy for the canonical lost sales inventory system with better performance than the optimal policy of the equivalent canonical back-order system. The proof uses a comparison based on associated random variables.

3. We prove that the projected inventory level policy is asymptotically optimal as the penalty for a lost sale approaches infinity under mild conditions on the demand distribution.
4. We prove that the projected inventory level policy is a 1-step policy improvement upon the constant order policy and dominates the performance of the constant order policy when demand has an exponential distribution.
5. We prove that the projected inventory level policy is asymptotically optimal as the lead time approaches infinity when demand has an exponential distribution. The proof approach is to show dominance of the PIL policy over the constant order policy such that it inherits asymptotic optimality properties.
6. We demonstrate numerically that the projected inventory level has superior performance also outside of the regimes where it is asymptotically optimal. We also provide results that aid with efficient computation within this class of policies.

## 2. Brief Literature Review

The canonical lost sales inventory system was first studied by Karlin and Scarf (1958) and found to have a more complicated optimal ordering policy than the canonical inventory system with back-orders. The optimal ordering policy for the lost sales system depends on each outstanding order. Sensitivities of the optimal ordering decision to each outstanding order were first characterized by Morton (1969) and the analysis was later streamlined by Zipkin (2008b). Despite these results, computation and implementation of the optimal policy is not practical. Most of the literature studies heuristic policies without any optimality guarantees (e.g. van Donselaar et al. 1996, Bijvank and Johansen 2012, Sun et al. 2014, van Jaarsveld 2020) but with numerically favorable performance. Notable exceptions are Levi et al. (2008) who prove that their dual balancing policy has an optimality gap of at most 100% and Chen et al. (2014) who provide a pseudo polynomial time approximation scheme. Another active area of research is the study of base-stock policies when the demand distribution is unknown and must be learned online (Huh et al. 2009a, Zhang et al. 2020, Agrawal and Jia 2022). We refer to Bijvank and Vis (2011) for a general review of lost sales inventory models.

We use the remainder of this brief literature review to focus on asymptotic optimality results as this is the focus of this paper. The notion of asymptotic optimality in lost sales and other difficult inventory systems has gained recent traction. Goldberg et al. (2021) provide a survey of such results and outline methodologies to prove such results. The most important results for the lost sales inventory system are asymptotic optimality of base-stock policies as the cost of losing a sale approaches infinity (Huh et al. 2009b, Bijvank et al. 2014) and asymptotic optimality of constant order policies as the lead-time approaches infinity (Goldberg et al. 2016, Xin and Goldberg 2016, Bu et al. 2020, Xin 2020, Bai et al. 2023). A natural question is whether any intuitive policies exist that are asymptotically optimal in both regimes. Xin (2020) propose the capped base-stock policy which was first introduced by Johansen and Thorstenson (2008). Under this policy orders are placed to reach a base-stock level except when this would cause the order size to exceed the cap. The parameters of this policy can be set such that it effectively reduces to either a constant order or base-stock policy. Xin (2020) show that such a policy is asymptotically optimal for long lead-times. It is straightforward to show that a capped base-stock policy is also asymptotically optimal in the other regime. The PIL policy has a single parameter that cannot be set such that it effectively reduces to either a constant order or base-stock policy. Despite this, we provide asymptotic optimality results both for long lead-times and high per unit lost sales costs.

Our analysis is based on comparison against simpler policies for which asymptotic optimality results have already been established. The comparison uses association of random variables when comparing the PIL with the base-stock policy. A similar technique has been used to bound the order fill-rates in assemble-to-order systems by Song (1998). We use policy improvement to compare the PIL policy to the constant order policy. This idea is often used to create an improved policy for a system that suffers from the curse of dimensionality so that only a simple policy can be analyzed (see e.g. Tijms 2003, Haijema et al. 2008). We are not aware of prior work that uses association of random variables and/or policy improvement to establish asymptotic optimality results.

### 3. Model

We consider a periodic review lost sales inventory system. In each period  $t \in \mathbb{N}_0$  (where  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ ) we place an order that will arrive after a lead time of  $\tau \in \mathbb{N}_0$  periods, i.e. at the start of period  $t + \tau$ . The inventory level at the beginning of period  $t$ , after receiving the order that was placed in period  $t - \tau$ , is denoted  $I_t$ . We denote the inventory level at the end of period  $t$  as  $J_t$ . The order placed in period  $t$  is denoted  $q_{t+\tau}$  such that the order that arrives in period  $t$  is denoted  $q_t$ . Demand in period  $t$  is denoted by  $D_t$  and  $\{D_t\}_{t=0}^\infty$  is an i.i.d. sequence of random variables with  $\mu := \mathbb{E}[D_t] < \infty$ . We let  $D$  denote the generic one period demand random variable and its distribution  $F(x) := P(D \leq x)$ .

Demand is satisfied from inventory whenever possible. If demand in a period exceeds the inventory level, there will be lost sales denoted by  $L_t$  in period  $t$ . The system dynamics are given as:

$$L_t = (D_t - I_t)^+ \quad (1)$$

$$J_t = (I_t - D_t)^+ \quad (2)$$

$$I_t = (I_{t-1} - D_{t-1})^+ + q_t = J_{t-1} + q_t, \quad (3)$$

where  $(x)^+ := \max(0, x)$ . The state of this system in period  $t$  is given by  $\mathbf{x}_t = (I_t, q_{t+1}, q_{t+2}, \dots, q_{t+\tau-1}) \in \mathbb{R}_+^\tau$ . Since we focus on long run costs, and for ease of exposition, we assume the initial state  $\mathbf{x}_0$  to be  $\mathbf{0}$ , i.e.  $I_0 = 0$  and  $q_t = 0$  for  $t < \tau$  unless stated otherwise (cf. Xin and Goldberg 2016). For convenience in notation we define  $D[a, b] := \sum_{i=a}^b D_i$  and similarly define  $I[a, b]$ ,  $J[a, b]$ ,  $q[a, b]$  and  $L[a, b]$ .

In each period  $t$  we may decide  $q_{t+\tau} \geq 0$  based on  $\mathbf{x}_t$  and  $t$ . For our purposes, it will be convenient and sufficient to represent policies by a set of functions  $A = \{A_t, t \in \mathbb{N}_0\}$ , where  $A_t$  maps states  $\mathbf{x}_t \in \mathbb{R}_+^\tau$  to actions  $q_{t+\tau} \geq 0$ .

We consider three policies in the analytical sections of this paper. The base-stock policy  $B^S$  (cf. Huh et al. 2009b) and the constant order policy  $C^r$  (cf. Xin and Goldberg 2016) are defined as follows:

$$B_t^S(\mathbf{x}_t) := (S - I_t - q[t+1, t+\tau-1])^+$$

$$C_t^r(\mathbf{x}_t) := r.$$

Here,  $S$  and  $r$  are the base-stock level and the constant order quantity, respectively. We assume for stability that  $r < \mu$  (cf. Xin and Goldberg 2016). The projected inventory level policy  $P^U$  with projected inventory level  $U$  is given by

$$P_t^U(\mathbf{x}_t) := (U - \mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t])^+ = (U - I_t - q[t+1, t+\tau-1] - \mathbb{E}[L[t, t+\tau-1]|\mathbf{x}_t] + \tau\mu)^+. \quad (4)$$

Note that the expectations in (4) take into account the state at time  $t$ . Therefore, at time  $t$  the PIL policy places an order to raise the expected inventory level at time  $t+\tau$  to  $U$ , if possible.

The following lemma specifies conditions that ensure that the projected inventory level  $U$  can be attained in every period; the proof is in Appendix EC.1.

LEMMA 1. *For any given PIL policy, it is possible to place a non-negative order in each period  $t \geq 0$  to attain the projected inventory level  $U \geq 0$  provided that it is possible to do so in period 0 (i.e. provided  $\mathbf{x}_0$  satisfies  $\mathbb{E}[J_{\tau-1}|\mathbf{x}_0] \leq U$ ).*

Any excess inventory at the end of a period incurs a holding cost  $h > 0$  per item per period. Any lost sales accrued during a period incur a lost sales penalty cost  $p > 0$  per item lost. Denote the costs incurred in period  $t$  by  $c_t := hJ_t + pL_t$ , and let  $c[a, b] := \sum_{t=a}^b c_t$ . We write  $c[a, b](A)$  to make the dependence on the policy  $A$  explicit. The cost-rate associated with a policy  $A$  is then

$$C(A) := \limsup_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T - \tau + 1} c[\tau, T](A) \right]$$

Let  $S^* \in \operatorname{argmin}_S C(B^S)$ ,  $r^* \in \operatorname{argmin}_r C(C^r)$ , and  $U^* \in \operatorname{argmin}_U C(P^U)$  denote the optimal base-stock level, constant order quantity, and projected inventory level respectively. Let  $C^*$  denote the long run expected costs of an optimal policy.

#### 4. Long lead time asymptotics

Constant-order policies were proven to be asymptotically optimal for long lead-time by Goldberg et al. (2016) (see also Xin and Goldberg 2016). This result has deepened the understanding of lost sales inventory systems. Empirically, we observe that the PIL policy outperforms the



constant order policy, also for long lead-times. In this sense, the PIL policy is unlike the base-stock policy, which cannot match the constant order policy performance for large lead-times. In this section, we will theoretically underpin this finding. In particular, we prove that the projected inventory policy is in expectation superior to the constant order policy when demand is exponentially distributed.

Our analysis is based on the following simple idea. Consider the total costs incurred from time  $t + \tau$  up to time  $T$ , given the state  $\mathbf{x}_t$  and order  $q$  and assuming  $q_k = r$  for  $k > t + \tau$ , and denote this total cost by  $F^T(q) = hJ[t + \tau, T] + pL[t + \tau, T]$ . Define  $f(q|\mathbf{x}_t) = \lim_{T \rightarrow \infty} \mathbb{E}[F^T(q) - F^T(r) | \mathbf{x}_t]$ . A sensible policy may decide  $q_{t+\tau} \in \operatorname{argmin}_{q \geq 0} f(q|\mathbf{x}_t)$  for any pipeline  $\mathbf{x}_t$ . This policy may be recognized as a single-step policy improvement to the constant order policy, and it may therefore be expected to dominate the constant order policy with order quantity  $r$ .

In the following, we sketch a heuristic argument that shows  $P^U(\mathbf{x}_t) \in \operatorname{argmin}_{q \geq 0} f(q|\mathbf{x}_t)$  with  $U = p(\mu - r)/h$ . This heuristic argument as well as the intuition that this policy dominates the constant order policy will be made rigorous in Section 4.1.

Let us determine the  $q = q_{t+\tau}$  that minimizes  $f$  for some initial state  $\mathbf{x}$ . Increasing  $q$  by  $\epsilon$  has two effects on the infinite horizon costs: 1)  $\epsilon$  more demand is eventually satisfied (not lost) and a penalty cost  $p\epsilon$  is averted 2) From time  $t + \tau$  until the first stockout, the inventory level increases by  $\epsilon$  (for  $\epsilon$  small). Suppose  $t + \tau = 0$  for notational convenience. Let  $R$  denote the time of the first lost sale from period 0:  $R = \min\{t \in \mathbb{N}_0 | L_t > 0\}$ . Then

$$\frac{df(q)}{dq} = \lim_{\epsilon \rightarrow 0} \frac{f(q + \epsilon) - f(q)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{-p\epsilon + h\mathbb{E}[R]\epsilon}{\epsilon} = h\mathbb{E}[R] - p. \quad (5)$$

The expectation of  $R$  can be found by noting that  $R$  is a stopping time and  $\mathbb{E}[L_R] = \mu$  due to the memoryless property of the exponential demand distribution. For  $t < R$  there are no stockouts so that  $J_t = J_{t-1} + r - D_t$ . Thus  $J_t = I_0 - D[0, t] + tr = J_{-1} + q - D[0, t] + tr$  while for  $t = R$  we have  $J_R = 0$  and  $L_R = -J_{-1} - q + D[0, R] - Rr$ . In particular this implies (using Wald's identity)  $\mathbb{E}[L_R] = -\mathbb{E}[J_{-1}] - q + (\mathbb{E}[R] + 1)\mu - \mathbb{E}[R]r$ . Using that also  $\mathbb{E}[L_R] = \mu$  and solving for  $\mathbb{E}[R]$  yields

$\mathbb{E}[R] = \frac{\mathbb{E}[J_{-1}] + q}{\mu - r}$ . Thus by (5),  $f$  must be a parabola in  $q$  and  $df(q)/dq = 0$  if and only if  $q = \frac{p(\mu - r)}{h} - \mathbb{E}[J_{-1}]$ . Since  $I_0 = J_{-1} + q$ , this holds if and only if  $q$  follows a projected inventory level policy with level  $U = \frac{p(\mu - r)}{h}$ .

#### 4.1. Dominance of PIL policies over COP policies

The heuristic argument above can be made rigorous as follows. Observe that the orders under  $C^r$  are independent of the state. Hence, the *bias* (or relative value function) of the  $C^r$  policy can be expressed as a function of inventory level only as in the following definition.

DEFINITION 1. Let  $r \in [0, \mathbb{E}[D]]$ , and let  $g^r = C(C^r)$  be the long run average costs of the  $C^r$  policy. Then the bias  $\mathcal{H}^r(\cdot) : \mathbb{R}^+ \rightarrow \mathbb{R}$  associated with  $C^r$  satisfies:

$$\mathcal{H}^r(x) = \mathbb{E}_D[h(x - D)^+ + p(D - x)^+ + \mathcal{H}^r((x - D)^+ + r)] - g^r, \quad (6)$$

for any non-negative  $x \geq 0$ . To make the bias unique we also impose  $\mathcal{H}^r(0) = 0$ .

This bias  $\mathcal{H}^r(x)$  can be interpreted as the additional cost over an infinite horizon of starting with  $x$  items in inventory and a pipeline with orders of size  $r$  instead of starting with 0 items in inventory and a pipeline with orders of size  $r$ . Intuitively,  $f(q|\mathbf{x}_t)$  equals  $\mathbb{E}[\mathcal{H}^r(J_{t+\tau-1} + q) - \mathcal{H}^r(J_{t+\tau-1} + r)|\mathbf{x}_t]$ , and informed by the heuristic argument made earlier one can guess that, like  $f$ , this bias should be a parabola.

LEMMA 2. *When demand has an exponential distribution, the bias of the constant order policy with constant order quantity  $r < \mu$  is a parabola:*

$$\mathcal{H}^r(x) = \frac{h}{2(\mu - r)}x^2 - px, \quad (7)$$

and  $g^r = C(C^r) = p(\mu - r) + h\frac{r^2}{2(\mu - r)}$ .

The proof of Lemma 2 can be found in Appendix EC.2 and is a straightforward verification that the proposed solution satisfies the definition.

Confirming our intuition, we next show that the policy found by a single improvement step on the bias of a constant order policy is a projected inventory level policy.

LEMMA 3. Let  $D$  be exponentially distributed and for any  $h, p, r < \mu$ , let  $U(r) = p(\mu - r)/h$ . For any  $\mathbf{x}_t$ , if  $q_{t+\tau} = P^{U(r)}(\mathbf{x}_t)$  then  $q_{t+\tau} \in \operatorname{argmin}_{q \geq 0} \mathbb{E}[\mathcal{H}^r(J_{t+\tau-1} + q) \mid \mathbf{x}_t]$ .

*Proof.* Using (7), one may derive  $\mathcal{H}^r(x) = a_1(x - U(r))^2 + a_2$  with  $a_1 = \frac{h}{2(\mu-r)} > 0$  and  $a_2 = \frac{-p^2(\mu-r)}{2h}$ , thus  $U(r)$  is the unique minimizer of  $\mathcal{H}^r(\cdot)$ . Now observe that

$$\begin{aligned} \mathbb{E}[\mathcal{H}^r(J_{t+\tau-1} + q) \mid \mathbf{x}_t] &= \mathbb{E}[a_1(J_{t+\tau-1} + q - U(r))^2 + a_2 \mid \mathbf{x}_t] \\ &= a_1 [\operatorname{Var}[J_{t+\tau-1} + q - U(r) \mid \mathbf{x}_t] + \mathbb{E}[J_{t+\tau-1} + q - U(r) \mid \mathbf{x}_t]^2] + a_2 \\ &= a_1 \operatorname{Var}[J_{t+\tau-1} \mid \mathbf{x}_t] + a_1(\mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t] + q - U(r))^2 + a_2, \end{aligned} \quad (8)$$

where the final equality follows because  $\operatorname{Var}[q \mid \mathbf{x}_t] = 0$  for any deterministic policy  $A$ . Clearly  $q = P^{U(r)}(\mathbf{x}_t) = (U(r) - \mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t])^+$  minimizes (8).  $\square$

For our continuous state-space, continuous action-space model, there appear to be no standard Markov decision process results that can be leveraged to prove from Lemma 3 that the PIL dominates the constant order policy. Our proof uses the following result, the proof of which appears in Appendix EC.3:

LEMMA 4. Let  $t_1 \leq t_2$ ,  $t_1, t_2 \in \mathbb{N}_0$ ,  $r \in [0, \mathbb{E}[D]]$ ,  $g^r = C(C^r)$ , and suppose  $q_t = r$  for all  $t \in \{t_1 + 1, \dots, t_2\}$ . Then

$$\mathbb{E}_{D_{t_1}, \dots, D_{t_2}} [c[t_1, t_2](C^r) \mid I_{t_1}] = \mathcal{H}^r(I_{t_1}) - \mathbb{E}_{D_{t_1}, \dots, D_{t_2}} [\mathcal{H}^r(I_{t_2+1}) \mid I_{t_1}] + (t_2 + 1 - t_1)g^r.$$

We are now ready to establish the main result of this section.

THEOREM 1. If demand has an exponential distribution then the best PIL policy  $P^{U^*}$  outperforms the best constant order policy  $C^{r^*}$ . In particular  $C(P^{U^*}) \leq C(P^{U(r^*)}) \leq C(C^{r^*})$  for any  $\tau \in \mathbb{N}_0$ , where  $U(r)$  is given by Lemma 3.

*Proof.* We focus on bounding  $\mathbb{E}[c[\tau, T](P^{U(r^*)}) - c[\tau, T](C^{r^*})]$ . A device in the proof will be a policy  $\mathcal{G}^{\tilde{t}}$  that places the first  $\tilde{t} \in \mathbb{N}_0$  orders using the projected inventory policy  $P^{U(r^*)}$ , and subsequent orders using the optimal constant-order policy  $C^{r^*}$ :

$$\mathcal{G}_t^{\tilde{t}}(\mathbf{x}) := \begin{cases} P_t^{U(r^*)}(\mathbf{x}), & t < \tilde{t} \\ C_t^{r^*}(\mathbf{x}) = r^*, & t \geq \tilde{t}. \end{cases}$$

Let  $I_t(A)$  denote the random variable  $I_t$  when policy A is adopted, and let  $\bar{t} = \tilde{t} + \tau$ . We will compare expected interval costs for the policies  $\mathcal{G}^{\tilde{t}+1}$  and  $\mathcal{G}^{\tilde{t}}$ :

$$\begin{aligned} \mathbb{E} \left[ c[\tau, T](\mathcal{G}^{\tilde{t}+1}) - c[\tau, T](\mathcal{G}^{\tilde{t}}) \right] &= \mathbb{E} \left[ c[\bar{t}, T](\mathcal{G}^{\tilde{t}+1}) - c[\bar{t}, T](\mathcal{G}^{\tilde{t}}) \right] \\ &= \mathbb{E} \left[ \mathbb{E}_{D_{\bar{t}}, \dots, D_T} \left[ c[\bar{t}, T] \middle| I_{\bar{t}}(\mathcal{G}^{\tilde{t}+1}) \right] - \mathbb{E}_{D_{\bar{t}}, \dots, D_T} \left[ c[\bar{t}, T] \middle| I_{\bar{t}}(\mathcal{G}^{\tilde{t}}) \right] \right] \\ &= \mathbb{E} [\mathcal{H}(I_{\bar{t}}(\mathcal{G}^{\tilde{t}+1})) - \mathcal{H}(I_{\bar{t}}(\mathcal{G}^{\tilde{t}})) - \mathcal{H}(I_{T+1}(\mathcal{G}^{\tilde{t}+1})) + \mathcal{H}(I_{T+1}(\mathcal{G}^{\tilde{t}}))]. \quad (9) \end{aligned}$$

Here and elsewhere in this proof,  $\mathcal{H}$  denotes  $\mathcal{H}^{r^*}$ . Also, the first equality follows because  $\mathcal{G}^{\tilde{t}+1}$  and  $\mathcal{G}^{\tilde{t}}$  coincide for  $t < \tilde{t}$ , and thus, since  $\mathbf{x}_0 = \mathbf{0}$  for  $t \leq \tilde{t}$  the distributions of  $\mathbf{x}_t$ ,  $J_{t+\tau-1}$ ,  $L_{t+\tau-1}$  and  $c_{t+\tau-1}$  are the same for the two policies. The second equality is by conditioning on the inventory at time  $\bar{t}$ . For the third equality, note that  $\mathcal{G}_t^{\tilde{t}+1} = \mathcal{G}_t^{\tilde{t}} = C^{r^*}$  for  $t \geq \tilde{t} + 1$ , hence  $q_t = r^*$  for  $t \geq \tilde{t} + 1 + \tau$  for both policies, and hence, we can substitute the identity of Lemma 4.

Now note that  $I_{\bar{t}}(\mathcal{G}^{\tilde{t}+1}) = J_{\bar{t}-1} + P_{\bar{t}}^{U(r^*)}(\mathbf{x}_{\bar{t}})$  and  $I_{\bar{t}}(\mathcal{G}^{\tilde{t}}) = J_{\bar{t}-1} + r^*$ , while  $\mathbf{x}_{\bar{t}}$  and  $J_{\bar{t}-1}$  are identically distributed for both policies since they coincide for  $t < \tilde{t}$ . We condition on  $\mathbf{x}_{\bar{t}}$  and find:

$$\begin{aligned} \mathbb{E} \left[ \mathcal{H} \left( I_{\bar{t}}(\mathcal{G}^{\tilde{t}+1}) \right) - \mathcal{H} \left( I_{\bar{t}}(\mathcal{G}^{\tilde{t}}) \right) \right] &= \mathbb{E} \left[ \mathbb{E} \left[ \mathcal{H} \left( I_{\bar{t}}(\mathcal{G}^{\tilde{t}+1}) \right) - \mathcal{H} \left( I_{\bar{t}}(\mathcal{G}^{\tilde{t}}) \right) \middle| \mathbf{x}_{\bar{t}} \right] \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \mathcal{H} \left( J_{\bar{t}-1} + P_{\bar{t}}^{U(r^*)}(\mathbf{x}_{\bar{t}}) \right) - \mathcal{H}(J_{\bar{t}-1} + r^*) \middle| \mathbf{x}_{\bar{t}} \right] \right] \\ &= \mathbb{E} \left[ \min_{q \geq 0} \left( \mathbb{E} [\mathcal{H}(J_{\bar{t}-1} + q) - \mathcal{H}(J_{\bar{t}-1} + r^*) \middle| \mathbf{x}_{\bar{t}}] \right) \right] \\ &= \mathbb{E} \left[ \min_{q \in \mathbb{R}} \left( \mathbb{E} [\mathcal{H}(J_{\bar{t}-1} + q) - \mathcal{H}(J_{\bar{t}-1} + r^*) \middle| \mathbf{x}_{\bar{t}}] \right) \right] \\ &= -\mathbb{E} \left[ \mathbb{E} \left[ a_1 \left( r^* - P_{\bar{t}}^{U(r^*)}(\mathbf{x}_{\bar{t}}) \right)^2 \middle| \mathbf{x}_{\bar{t}} \right] \right] \\ &= -a_1 \mathbb{E} \left[ \left( r^* - P_{\bar{t}}^{U(r^*)}(\mathbf{x}_{\bar{t}}) \right)^2 \right] \quad (10) \end{aligned}$$

For the third equality, we use Lemma 3. For the fourth equality, observe that  $\mathbf{x}_0 = \mathbf{0}$ , and hence  $\mathbb{E}[J_{\bar{t}-1} | \mathbf{x}_{\bar{t}}] \leq U(r^*)$  (cf. Lemma 1), which implies that the minimum over  $q \in \mathbb{R}$  is attained by an element  $q \geq 0$ . For the fifth equality, we substitute equation (8) and cancel terms.

With this, we obtain:

$$\mathbb{E} [c[\tau, T](P^{U(r^*)}) - c[\tau, T](C^{r^*})] = \mathbb{E} [c[\tau, T](\mathcal{G}^{T+1-\tau}) - c[\tau, T](\mathcal{G}^0)]$$

$$\begin{aligned}
 &= \sum_{\tilde{t}=0}^{T-\tau} \mathbb{E} [c[\tau, T](\mathcal{G}^{\tilde{t}+1}) - c[\tau, T](\mathcal{G}^{\tilde{t}})] \\
 &= \sum_{\tilde{t}=0}^{T-\tau} \mathbb{E} [\mathcal{H}(I_{\tilde{t}}(\mathcal{G}^{\tilde{t}+1})) - \mathcal{H}(I_{\tilde{t}}(\mathcal{G}^{\tilde{t}})) - \mathcal{H}(I_{T+1}(\mathcal{G}^{\tilde{t}+1})) + \mathcal{H}(I_{T+1}(\mathcal{G}^{\tilde{t}}))] \\
 &= \mathbb{E}[I_{T+1}(\mathcal{G}^0)] - \mathbb{E}[I_{T+1}(\mathcal{G}^{T+1-\tau})] + \sum_{\tilde{t}=0}^{T-\tau} \mathbb{E} [\mathcal{H}(I_{\tilde{t}}(\mathcal{G}^{\tilde{t}+1})) - \mathcal{H}(I_{\tilde{t}}(\mathcal{G}^{\tilde{t}}))] \\
 &= \mathbb{E}[I_{T+1}(C^{r^*})] - \mathbb{E}[I_{T+1}(P^{U(r^*)})] - a_1 \sum_{\tilde{t}=0}^{T-\tau} \mathbb{E} \left[ \left( r^* - P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}}) \right)^2 \right]
 \end{aligned} \tag{11}$$

Here, the first equality holds by definition of  $\mathcal{G}^{\tilde{t}}$ . The second equality follows by expressing the difference as a telescoping sum. The third equality uses (9). For the fourth equality, we rearrange and cancel terms. The final equality holds by definition of  $\mathcal{G}^{\tilde{t}}$ , and by (10).

Using the definition of the cost-rate, we find

$$\begin{aligned}
 C(C^{r^*}) &= \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} c[\tau, T](C^{r^*}) \right] \\
 &= \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} \left( c[\tau, T](P^{U(r^*)}) + a_1 \sum_{\tilde{t}=0}^{T-\tau} \left( r^* - P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}}) \right)^2 \right) \right] \\
 &= \limsup_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} c[\tau, T](P^{U(r^*)}) \right] + a_1 \liminf_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} \sum_{\tilde{t}=0}^{T-\tau} \left( r^* - P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}}) \right)^2 \right]
 \end{aligned}$$

The first equality uses that for the constant order policy, the sequence converges such that the limit superior equals the limit. For the second equality, we use (11) and  $\lim_{T \rightarrow \infty} \frac{1}{T-\tau+1} (\mathbb{E}[I_{T+1}(C^{r^*})] - \mathbb{E}[I_{T+1}(P^{U(r^*)})]) = 0$ . That this limit goes to zero follows because the steady state inventory level under the constant order policy exists (with finite mean), and because the inventory level under the PIL policy is bounded from below by 0 and from above by  $U(r^*) + \tau\mu$ . This upper bound holds because  $P_t^{U(r^*)}(\mathbf{x}_t) \leq (U(r^*) + \tau\mu - I_t - q[t+1, t+\tau-1])^+$ , i.e. the inventory position after placing an order  $I_t + q[t+1, t+\tau-1] + q_\tau$  cannot rise above  $U(r^*) + \tau\mu$ , (and it is below this bound initially since  $\mathbf{x}_0 = \mathbf{0}$  by assumption). Thus the inventory level is bounded by  $U(r^*) + \tau\mu$ .

The third equality now follows because all limit points of the sequence  $\left( \mathbb{E} \left[ \frac{1}{T-\tau+1} \sum_{\tilde{t}=0}^{T-\tau} \left( r^* - P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}}) \right)^2 \right] \right)_{T=\tau, \tau+1, \dots}$  must be finite since  $P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}})$  is bounded below

by 0 and above by  $U(r^*) + \tau\mu$ . Now note that for each limit point  $x \in \mathbb{R}$  of that sequence, there must exist a subsequence that converges to that limit point. By the second equality above, the corresponding subsequence of  $\left(\mathbb{E}\left[\frac{1}{T-\tau+1}c[\tau, T](P^{U(r^*)})\right]\right)_{T=\tau, \tau+1, \dots}$  must converge to  $y = C(C^{r^*}) - x$ . Thus every limit point  $x$  of the former sequence must correspond to a limit point  $y = C(C^{r^*}) - x$  of the latter sequence, and vice versa. Also, the smallest limit point of the latter sequence must correspond to the largest limit point of the former sequence, which implies the third equality.

We thus find:

$$C(C^{r^*}) - C(P^{U(r^*)}) = a_1 \liminf_{T \rightarrow \infty} \frac{1}{T - \tau + 1} \sum_{\tilde{t}=0}^{T-\tau} \mathbb{E} \left[ \left( r^* - P_{\tilde{t}}^{U(r^*)}(\mathbf{x}_{\tilde{t}}) \right)^2 \right] \geq 0. \quad (12)$$

This completes the proof.  $\square$

It is noteworthy that the difference between the cost of a PIL policy and a constant order policy can be expressed as a function of the quadratic differences between the order decisions of both policies; see (12). The amounts by which the decisions of the PIL policy differ from a constant order policy also express how much better it performs.

#### 4.2. Asymptotic optimality as $\tau \rightarrow \infty$

With the results of Xin and Goldberg (2016), Theorem 1 establishes asymptotic optimality of the PIL policy as  $\tau$  grows large:

**THEOREM 2.** *The PIL policy is asymptotically optimal for long lead-times when demand has an exponential distribution:*

$$\lim_{\tau \rightarrow \infty} \left( C(P^{U(r^*)}) - C^* \right) = 0.$$

This result follows directly from our Theorem 1, and Theorem 1 in Xin and Goldberg (2016). In fact, it follows from these theorems that the optimality gap of PIL policy decays exponentially in  $\tau$ .

## 5. Penalty cost asymptotics

The performance of the PIL-policy in the asymptotic regime that  $p \rightarrow \infty$  will be studied by using bounds in terms of the related inventory system in which demand in excess of inventory is back-ordered rather than lost. Therefore, in this section, we first make a comparison to the related canonical back-order system in Section 5.1 and show that the cost of a PIL policy for the lost sales system has lower cost than the optimal policy for the canonical back-order system. Then we provide our main result that the PIL policy is asymptotically optimal as the cost of losing a sale approaches infinity in Section 5.2.

### 5.1. Comparison to back-order system

The back-order system with lead-time  $\tau \in \mathbb{N}_0$ , holding cost parameter  $h > 0$  per period per item, and back-order cost  $p > 0$  per period per item is much better understood than the same system with lost sales. Comparison between these two systems has been studied before by e.g. Janakiraman et al. (2007), Huh et al. (2009b), and Bijvank et al. (2014) and we mostly follow their notational conventions. The dynamics for the back-order system are:

$$I_t^{\mathcal{B}} = I_{t-1}^{\mathcal{B}} - D_{t-1} + q_t^{\mathcal{B}}, \quad J_t^{\mathcal{B}} = (I_t^{\mathcal{B}} - D_t)^+, \quad B_t = (D_t - I_t^{\mathcal{B}})^+, \quad (13)$$

where  $B_t$  and  $I_t^{\mathcal{B}}$  denote respectively the number of items on back-order and the inventory level in period  $t$ . We use the superscript  $\mathcal{B}$  to denote that quantities belong to the back-order system (rather than the lost sales system) and generally denote this system as  $\mathcal{B}$ . It is well known that the optimal policy for  $\mathcal{B}$  is a base-stock policy. Under this policy the order in each period  $t$  is placed to raise the inventory position to a fixed base-stock level  $S$ :

$$q_{t+\tau}^{\mathcal{B}} = S - I_t^{\mathcal{B}} - q^{\mathcal{B}}[t+1, t+\tau-1], \quad (14)$$

and the optimal base-stock level for system  $\mathcal{B}$  is given by the newsvendor equation:

$$S^* = \inf \left\{ S : \mathbb{P}(D[0, \tau] \leq S) \geq \frac{p}{p+h} \right\}. \quad (15)$$

The optimal average cost-rate for this system satisfies:

$$C^{\mathcal{B}*} = h\mathbb{E}[(S^* - D[0, \tau])^+] + p\mathbb{E}[(D[0, \tau] - S^*)^+]. \quad (16)$$

For the lost sales system we similarly define  $c_t^{\mathcal{B}} := hJ_t^{\mathcal{B}} + pB_t$  and  $c^{\mathcal{B}}[a, b] := \sum_{t=a}^b c_t^{\mathcal{B}}$ . We also write  $c^{\mathcal{B}}[a, b](A)$  to make the dependence on the control policy  $A$  explicit. With this notation we can express the cost-rate of a policy  $A$  in system  $\mathcal{B}$  as  $C^{\mathcal{B}}(A) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} c^{\mathcal{B}}[\tau, T](A) \right]$ . The lost sales system described in Section 3 is denoted by  $\mathcal{L}$ , and the optimal cost for this system is still denoted by  $C^*$ . Note that both systems are defined on the same probability space induced by the initial state and demand sequence. The main result of Janakiraman et al. (2007) is that  $C^* \leq C^{\mathcal{B}*}$ , which is established via an ingenious stochastic comparison technique. The main result of this section is that the best PIL-policy for  $\mathcal{L}$  achieves lower cost than the optimal policy for  $\mathcal{B}$ :  $C(P^{U*}) \leq C^{\mathcal{B}*}$ . Via  $C^* \leq C(P^{U*}) \leq C^{\mathcal{B}*}$ , this result constitutes the first *constructive* proof of the main result of Janakiraman et al. (2007): Unlike their stochastic comparison proof, we identify a specific (PIL) policy that yields a cost-rate for system  $\mathcal{L}$  that is lower than the optimal cost-rate for system  $\mathcal{B}$ . In addition to this, this result also enables us to leverage results in Huh et al. (2009b) to show (under mild conditions) that the PIL-policy is asymptotically optimal as  $p$  grows large.

The main idea behind the proof of this result is that the base-stock policy with base-stock level  $S > \tau\mu$  in  $\mathcal{B}$  is also a PIL-policy with projected inventory  $S - \tau\mu$  in system  $\mathcal{L}$ . Indeed observe that

$$I_{t+\tau}^{\mathcal{B}} = I_t^{\mathcal{B}} + q[t+1, t+\tau] - D[t, t+\tau-1] = S - D[t, t+\tau-1],$$

so that  $\mathbb{E}[I_{t+\tau}^{\mathcal{B}}] = S - \tau\mu$ . From this it immediately follows that  $C^{\mathcal{B}}(P^{S^* - \tau\mu}) = C^{\mathcal{B}}(B^{S^*}) = C^{\mathcal{B}*}$  when  $S^* \geq \tau\mu$ .

We will first show that  $\mathbb{E}[L_t] \leq \mathbb{E}[B_t]$  for any period  $t \geq \tau$  when  $\mathcal{L}$  and  $\mathcal{B}$  operate under the same PIL-policy with level  $U \geq 0$ . From that, we conclude that the cost-rate of  $\mathcal{L}$  under the optimal PIL policy is smaller than the optimal cost-rate for system  $\mathcal{B}$ .

The following technical lemma is needed in subsequent results. Its proof is in Appendix EC.4.



LEMMA 5. Let  $X$  and  $Y$  be random variables with joint distribution function  $F(x, y)$ ,  $-\infty < x, y < \infty$ . Then  $\mathbb{E}[(X + Y)^+] = \int_{-\infty}^{\infty} \mathbb{P}(X \geq z, Y \geq -z) dz$ .

Define the random variables

$$Y = L[0, \tau - 1] - \mathbb{E}[L[0, \tau - 1]], \quad \text{and} \quad X = D_\tau - I_\tau = D[0, \tau] - S - Y,$$

for  $S \geq \tau\mu$ . Observe that  $X^+ = L_\tau$  and that  $X + Y = D[0, \tau] - S$  such that  $(X + Y)^+ = B_\tau$  under a PIL-policy with level  $U = S - \tau\mu \geq 0$ . Our aim will be to prove that  $\mathbb{E}[L_\tau] \leq \mathbb{E}[B_\tau]$ . To this end, we first prove that  $X$  and  $Y$  are *associated* random variables (cf. Esary et al. (1967)).

DEFINITION 2. The random variables  $(A_1, \dots, A_n) = \mathbf{A}$  are said to be associated if

$$\text{Cov}[f(\mathbf{A}), g(\mathbf{A})] \geq 0$$

for all non-decreasing functions  $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$  for which the covariance above exists.

LEMMA 6.  $X$  and  $Y$  are associated random variables.

*Proof.* The random variables  $\mathbf{D} = (D_0, \dots, D_\tau)$  are associated by Theorem 2.1 of Esary et al. (1967). By property  $P_4$  of Esary et al. (1967) it suffices to show that  $Y = f(\mathbf{D})$  and  $X = g(\mathbf{D})$  are non-decreasing functions (element wise).

We derive an expression for  $L[0, t]$ : The cumulative demands lost until  $t$ . Each arriving demand  $D[0, t]$  until  $t$  is either satisfied from inventory or lost. When  $L_t > 0$ , the cumulative amount satisfied must equal the cumulative available inventory  $I_0 + q[1, t]$ , hence  $D[0, t] = L[0, t] + (I_0 + q[1, t]) \Rightarrow L[0, t] = D[0, t] - I_0 - q[1, t]$ . When  $L_t = 0$ , we have  $L[0, t] = L[0, t - 1]$ , and hence in general we find:

$$L[0, t] = \max_{k \in \{0, \dots, t\}} (D[0, k] - I_0 - q[1, k])^+. \quad (17)$$

(To see this, note that the maximum is attained for the  $k$  which corresponds to the period with the last stockout until  $t$ .) Next for  $Y = f(\mathbf{D})$  we have

$$Y = L[0, \tau - 1] - \mathbb{E}[L[0, \tau - 1]] = \max_{k \in \{0, \dots, \tau - 1\}} (D[0, k] - I_0 - q[1, k])^+ - \mathbb{E}[L[0, \tau - 1]], \quad (18)$$

which is clearly non-decreasing in each  $D_i$ ,  $i \in \{1, \dots, \tau\}$ . (Note that  $Y$  is independent of, and thus non-decreasing in  $D_\tau$ .) Finally observe that

$$X = g(\mathbf{D}) = D[0, \tau] - S - \max_{k \in \{0, \dots, \tau-1\}} (D[0, k] - I_0 - q[1, k])^+ + \mathbb{E}[L[0, \tau-1]]. \quad (19)$$

Note that  $dX/dD_\tau = 1$  and  $dX/dD_i \in \{0, 1\}$  for  $i \in \{0, \dots, \tau-1\}$ , thus  $g$  is non-decreasing.  $\square$

LEMMA 7. *If system  $\mathcal{B}$  and  $\mathcal{L}$  both operate under a PIL policy with level  $U \geq 0$ , then  $\mathbb{E}[B_t] \geq \mathbb{E}[L_t]$  for any period  $t \geq \tau$ .*

*Proof.* We will show that  $\mathbb{E}[B_\tau] \geq \mathbb{E}[L_\tau]$  using only that  $q_\tau$  can be placed to attain  $U \geq 0$ . By Lemma 1 this implies the result.

Let  $\tilde{X}$  and  $\tilde{Y}$  be two *independent* random variables with the same marginal distribution as  $X$  and  $Y$  respectively. Observe that

$$\begin{aligned} \mathbb{E}[B_\tau] &= \mathbb{E}[(X + Y)^+] \\ &= \int_{z=-\infty}^{\infty} \mathbb{P}(X \geq z, Y \geq -z) dz \end{aligned} \quad (20)$$

$$\begin{aligned} &\geq \int_{z=-\infty}^{\infty} \mathbb{P}(X \geq z) \mathbb{P}(Y \geq -z) dz \\ &= \int_{z=-\infty}^{\infty} \mathbb{P}(\tilde{X} \geq z) \mathbb{P}(\tilde{Y} \geq -z) dz = \mathbb{E}[(\tilde{X} + \tilde{Y})^+], \end{aligned} \quad (21)$$

where (20) follows from Lemma 5 and (21) follows from Theorem 5.1 of Esary et al. (1967). Now continuing and using that  $\tilde{X}$  and  $\tilde{Y}$  are independent, we find

$$\mathbb{E}[(\tilde{X} + \tilde{Y})^+] \geq \mathbb{E}[(\tilde{X} + \mathbb{E}[\tilde{Y}])^+] = \mathbb{E}[\tilde{X}^+] \quad (22)$$

where the inequality follows from Jensen's inequality and the independence between  $\tilde{X}$  and  $\tilde{Y}$ , and the equality holds since  $\mathbb{E}[\tilde{Y}] = 0$ . Note  $\mathbb{E}[\tilde{X}^+] = \mathbb{E}[X^+] = \mathbb{E}[L_t]$ .  $\square$

LEMMA 8. *If system  $\mathcal{B}$  and  $\mathcal{L}$  both operate under a PIL policy with level  $U \geq 0$ , then for any initial state  $\mathbf{x}$  such that projected inventory  $U \geq 0$  can be attained, we have  $p\mathbb{E}[L_t] + h\mathbb{E}[J_t] = \mathbb{E}[c_t] \leq \mathbb{E}[c_t^{\mathcal{B}}] = p\mathbb{E}[B_t] + h\mathbb{E}[J_t^{\mathcal{B}}]$  for any  $t \geq \tau$ .*

*Proof.* We will show that  $\mathbb{E}[c_\tau] \leq \mathbb{E}[c_\tau^{\mathcal{B}}]$  using only that  $q_\tau$  can be placed to attain  $U \geq 0$ . By Lemma 1 this implies the result. The inventory level in  $\mathcal{L}$  at the time of arrival of order  $q_\tau$  is given by:

$$I_\tau = I_0 + q[1, \tau] - D[0, \tau - 1] + L[0, \tau - 1]. \quad (23)$$

Under a PIL-policy with projected inventory  $S - \tau\mathbb{E}[D]$ , system  $\mathcal{L}$  will choose  $q_\tau$  such that  $\mathbb{E}[I_\tau] = S - \tau\mathbb{E}[D]$ . Using (23) and solving for  $q_\tau$  yields that

$$q_\tau = S - I_0 - q[1, \tau - 1] - \mathbb{E}[L[0, \tau - 1]]. \quad (24)$$

Substituting (24) back into (23) yields

$$I_\tau = S - D[0, \tau - 1] + L[0, \tau - 1] - \mathbb{E}[L[0, \tau - 1]]. \quad (25)$$

Now we have for the expected costs that will be incurred in period  $\tau$  by system  $\mathcal{L}$ :

$$\begin{aligned} \mathbb{E}[c_\tau] &= h\mathbb{E}[(I_\tau - D_\tau)^+] + p\mathbb{E}[(D_\tau - I_\tau)^+] \\ &= h\mathbb{E}[I_\tau - D_\tau] + h\mathbb{E}[(D_\tau - I_\tau)^+] + p\mathbb{E}[(D_\tau - I_\tau)^+] \\ &= h\mathbb{E}[S - D[0, \tau] + L[0, \tau - 1] - \mathbb{E}(L[0, \tau - 1])] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\ &= h\mathbb{E}[S - D[0, \tau]] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\ &= h\mathbb{E}\left[(S - D[0, \tau])^+\right] - h\mathbb{E}\left[(D[0, \tau] - S)^+\right] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\ &\leq h\mathbb{E}\left[(S - D[0, \tau])^+\right] + p\mathbb{E}\left[(D[0, \tau] - S)^+\right] = \mathbb{E}[c_\tau^{\mathcal{B}}] = C^{\mathcal{B}*}. \end{aligned} \quad (26)$$

The second and fourth equality follow from using the identity  $x = x^+ + (-x)^+$ , and the inequality follows from applying Lemma 7 twice.  $\square$

**THEOREM 3.** *If system  $\mathcal{B}$  is controlled by the optimal base-stock policy, or equivalently by a PIL policy with parameter  $S^* - \tau\mu$ , and if  $\mathcal{L}$  is controlled by a PIL-policy with PIL-level  $(S^* - \tau\mu)^+$ , then the cost-rate of system  $\mathcal{L}$  dominates the optimal cost-rate of system  $\mathcal{B}$ , that is*

$$C(\mathbf{P}^{U^*}) \leq C(\mathbf{P}^{(S^* - \tau\mu)^+}) \leq C^{\mathcal{B}*}.$$

*Proof.* Since  $\mathbf{x}_0 = \mathbf{0}$ , we can attain the projected inventory level  $U$  in  $\mathcal{L}$  in every period. (Regardless of this assumption, the number of periods for which it is not possible to attain a projected inventory level  $U$  in  $\mathcal{L}$  is finite almost surely.) Now consider two cases:  $S^* \geq \tau\mu$  and  $S^* < \tau\mu$ .

1. Case  $S^* \geq \tau\mu$ : Without loss of generality let period 0 be the first period in which it is possible to place an order to attain  $U = S^* - \tau\mu \geq 0$ . By Lemma 8 it holds that  $\mathbb{E}[c_t] \leq \mathbb{E}[c_t^{\mathcal{B}}] = C^{\mathcal{B}*}$  for all  $t \geq \tau$  which implies that  $C(P^{S^* - \tau\mu}) \leq C^{\mathcal{B}*}$ .
2. Case  $S^* < \tau\mu$ : Observe first that  $C(P^0) = p\mu$  because under  $U = 0$  there is no inventory and all demand is lost. We now have  $C^{\mathcal{B}*} = C^{\mathcal{B}}(P^{S^* - \tau\mu}) = p\mathbb{E}[(D[0, \tau] - S^*)^+] + h\mathbb{E}[(S^* - D[0, \tau])^+] \geq p\mathbb{E}[D[0, \tau] - S^*] > p\mu = C(P^0)$ , where the strict inequality holds because  $0 \leq S^* < \tau\mu$ .

That  $C(P^{U^*}) \leq C(P^{(S^* - \tau\mu)^+})$  follows from the definition of  $U^*$ .  $\square$

## 5.2. Asymptotic optimality as $p \rightarrow \infty$

To describe penalty cost asymptotics we need the following assumption on the distribution of lead time demand which is identical to assumption 1 of Huh et al. (2009b) and Bijvank et al. (2014):

ASSUMPTION 1. *The random variable  $D[0, \tau]$  has finite mean and is (i) bounded or (ii) is unbounded and  $\lim_{x \rightarrow \infty} \mathbb{E}[D[0, \tau] - x | D[0, \tau] > x] / x = 0$ .*

Assumption 1 is discussed in some detail in Section 3 of Huh et al. (2009b). All distributions commonly used to model demand, including Gaussian, gamma, Poisson, negative-binomial, Mixed Erlang, and Weibull distributions, satisfy this assumption.

THEOREM 4. *Under assumption 1, the best PIL-policy is asymptotically optimal for the lost sales inventory system as the cost of a lost sale increases:*

$$\lim_{p \rightarrow \infty} \frac{C(P^{(S^*(p) - \tau\mu)^+})}{C^*} = \lim_{p \rightarrow \infty} \frac{C(P^{U^*})}{C^*} = 1.$$

*Proof.* By Theorem 3 of Huh et al. (2009b) we have  $\lim_{p \rightarrow \infty} C^{\mathcal{B}*} / C^* = 1$ . Combining this with Theorem 3 yields the result.  $\square$

Recall that by Lemma 3, a 1-step policy improvement of a constant order policy yields a PIL policy. The former class of policies is not asymptotically optimal for  $p \rightarrow \infty$  and tends to perform poorly in this regime, while Theorem 4 demonstrates that the latter class of policies are asymptotically optimal as  $p \rightarrow \infty$ . Thus policy improvement qualitatively alters asymptotic performance.

## 6. Computational aspects

In this section we discuss the computation of the projected inventory level  $\mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t]$ , the optimization of the projected inventory level  $U$ , and efficient simulation estimators.

### 6.1. Inventory Projection

Implementation of the PIL policy requires that the projected inventory level,  $\mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t]$ , is computed every period  $t$ . This can be done relatively straightforwardly when demand has a discrete distribution by using the recursive expressions in (1) through (3). In this sub-section, we show how  $\mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t]$  can be computed similarly straightforwardly when demand has a Mixed Erlang (ME) distribution. The class of ME distributions is a powerful modeling tool because it can approximate any non-negative distribution arbitrarily closely (cf. Theorem 5.5.1 of Tijms (2003)), can be fitted easily on moments, and has many computational advantages in multi-echelon inventory theory (van Houtum 2006). We next define ME distributions and discuss how to project inventory levels when demand has a ME distribution.

Let  $\{E_{i,t}\}_{i=1}^{\infty}$  be an i.i.d. sequence of exponential random variables with mean  $1/\lambda$  for each  $t \in \mathbb{N}_0$ . Furthermore let  $\{K_t\}_{t=0}^{\infty}$  be a sequence of i.i.d. random variable on the non-negative integers with probability mass function  $\mathbb{P}(K_t = k) = \theta_k$ ,  $k \in \mathbb{N}_0$ . Demand has an ME distribution when  $D_t = \sum_{i=1}^{K_t} E_{i,t}$ . A common parameterization of mixed Erlang distributions is based on two moment fitting (e.g. Tijms 2003, van Houtum 2006). For convenience we provide this fitting procedure in Section EC.7.

In order to intuitively explain the efficient inventory projection method that follows, we imbue ME distributed demand with the following interpretation: Each period  $K_t$  customers arrive, each

demanding an exponentially distributed amount of stock. Recall that  $\mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t] = I_t + q[t+1, t+\tau-1] - \tau\mu + \mathbb{E}[L[t, t+\tau-1]|\mathbf{x}_t]$  so that in order to project the inventory level, it suffices to evaluate  $\mathbb{E}[L[t, t+\tau-1]|\mathbf{x}_t]$ . The main idea now is to count inventory and lost sales in terms of the number of customers (each with one exponential phase of demand) that can be satisfied. Let  $\tilde{I}_t$  denote the number of customers whose demand can be met fully with the inventory available at the beginning of period  $t$ . Then conditional on  $\mathbf{x}_t = (I_t, q_{t+1}, q_{t+2}, \dots, q_{t+\tau-1})$ ,  $\tilde{I}_t$  is Poisson distributed with mean  $\lambda I_t$ . Since  $K_t$  customers will arrive in period  $t$ , there will be  $\tilde{L}_t := (K_t - \tilde{I}_t)^+$  customers whose demand cannot be filled completely; note that for one of those customers the demand can be filled partially. Next, we crucially observe that for the customer whose demand is met partially, the amount of demand (in original units) that remains unfulfilled has an exponential distribution with mean  $1/\lambda$  due to the lack of memory of the exponential distribution. Thus  $\mathbb{E}[L_t|\mathbf{x}_t] = \lambda^{-1}\mathbb{E}[\tilde{L}_t|\mathbf{x}_t]$ . Similarly, there is inventory to satisfy the demand of another  $\tilde{J}_t := (\tilde{I}_t - K_t)^+$  customers at the end of period  $t$ . Let  $Q_{t+1}$  have a Poisson distribution with mean  $\lambda q_{t+1}$ . Then in period  $t+1$  there is inventory to satisfy the demand of  $\tilde{I}_{t+1} = \tilde{J}_t + Q_{t+1}$  customers. This reasoning can be continued to obtain the following dynamics:

$$\tilde{J}_t = (\tilde{I}_t - K_t)^+, \quad (27)$$

$$\tilde{L}_t = (K_t - \tilde{I}_t)^+, \quad (28)$$

$$\tilde{I}_{t+1} = \tilde{J}_t + Q_{t+1}, \quad (29)$$

with the initial conditions that  $\tilde{I}_t$  has a Poisson distribution with mean  $\lambda I_t$  and  $Q_t$  has a Poisson distribution with mean  $\lambda q_t$ , where  $1/\lambda$  is the mean of the exponential distributions associated with the ME demand. Note that since  $\tilde{I}_t$ ,  $\tilde{J}_t$  and  $\tilde{L}_t$  all denote a number of customers (whose demand can be met in full or is partially lost) they are distributed on the non-negative integers. Thus the distributions of  $\tilde{I}_{t+j}|\mathbf{x}_t$ ,  $\tilde{J}_t + j|\mathbf{x}_t$  and  $\tilde{L}_{t+j}|\mathbf{x}_t$  can be computed recursively for  $j \in \{0, \dots, \tau-1\}$ ; see Appendix EC.6. This recursion will generally require that the distributions of  $\tilde{I}_t$ ,  $\tilde{J}_t$ , and  $Q_t$  are truncated at a sufficiently high level.  $K_t$  will usually have a finite support  $k_{\max} = \sup\{k \in \mathbb{N} |$

$\theta_k > 0\}$ . In particular, under two moment fits,  $K_t$  will have a two-point distribution, see Section EC.7. Whenever  $k_{\max} < \infty$ ,  $\tilde{L}_{t+j}$  will be supported up to  $(j+1)k_{\max}$ . From this it follows that the distribution of  $\tilde{L}_{t+j}$  can be determined when the distributions of  $\tilde{I}_t$  and  $\tilde{J}_t$  are known up till  $(\tau+1)k_{\max}$ . Thus, the distributions of  $\tilde{I}_t$  and  $\tilde{J}_t$  can be truncated at  $(\tau+1)k_{\max}$  without affecting the distribution of  $\tilde{L}_{t+j}$ . This is convenient for computational purposes as there is no need to truncate any infinite summations. Thus we can compute the projected inventory level as

$$\mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t] = I_t + q[t+1, t+\tau-1] - \tau\mu + \frac{1}{\lambda} \mathbb{E} \left[ \tilde{L}[t, t+\tau-1] \mid \mathbf{x}_t \right], \quad (30)$$

when demand has a ME distribution. The theoretical running time complexity of this procedure is polynomial in  $k_{\max}$  and  $\tau$ . In practice, the procedure is efficient (see Section 7.3).

## 6.2. Optimization of the projected inventory level policy

We next discuss the optimization of  $U$ , i.e., we discuss and study the problem:

$$\min_{U \geq 0} C(P^U)$$

The expected cost incurred in period  $t$  can be written as  $c_t = h(I_t - D_t)^+ + p(D_t - I_t)^+ = h(I_t - D_t) + (h+p)(D_t - I_t)^+$ . Since  $\mathbb{E}[I_t] = U$  by Lemma 1, we obtain

$$C(P^U) = hU - h\mathbb{E}[D] + (h+p) \limsup_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} L[\tau, T](P^U) \right] \quad (31)$$

The following theorem will facilitate the optimization problem. Its proof is in Appendix EC.5.

LEMMA 9. *For any given  $\mathbf{x}_0$  and demand sequence  $D_0, \dots, D_{t-1}, D_t, \dots, D_{t+\tau}$ :*

1. *The cumulative inventory ordered until period  $t+\tau$  (i.e.  $q[1, t+\tau](P^U) \mid D_0, \dots, D_{t-1}$ ) is non-decreasing and concave in  $U$ .*
2. *The cumulative lost demand until period  $t+\tau$  (i.e.  $L[1, t+\tau](P^U) \mid D_0, \dots, D_{t+\tau}$ ) is non-increasing and convex in  $U$ .*

The next theorem follows immediately from Lemma 9 and facilitates optimization of the PIL-level.

THEOREM 5.  $C(P^U)$  is convex in  $U$ .

*Proof.*  $\limsup_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T-\tau+1} L[\tau, T](P^U) \right]$  is non-increasing and convex in  $U$  by Lemma 9. Combining with (31) yields the result.  $\square$

### 6.3. Low variance simulation estimator

The performance of the PIL policy can only be evaluated through simulation. The obvious simulation estimate of  $C(P^U)$  (or the cost-rate of any other policy) is  $\frac{1}{t_2-t_1} \sum_{t=t_1+1}^{t_2} (pL_t + hJ_t)$ , where  $t_2 - t_1$  is a simulation run-length sufficiently long to obtain an accurate estimate and  $t_1$  is the warm-up period. For the projected inventory level policy we can exploit results from Sections 6.2 and 6.1 to determine an unbiased estimator with much lower variance than this obvious choice. This variance reduction allows us to obtain accurate simulation estimates with much smaller simulation run-times. First we observe from equation (31) that another unbiased estimator of  $C(P^U)$  is given by  $\frac{1}{t_2-t_1} \sum_{t=t_1+1}^{t_2} ((p+h)L_t + h(U-\mu))$ . Next we observe that in this estimator,  $L_t$ , can be replaced by  $\mathbb{E}[L_{t+\tau-1} | \mathbf{x}_t]$ . This last expectation has a much smaller variance than  $L_t$  and its value is computed in each period already as part of projecting the inventory level; see (30). Thus an unbiased estimator for  $C(P^U)$  is given by

$$\widehat{C(P^U)} = \frac{1}{t_2-t_1} \sum_{t=t_1+1}^{t_2} ((p+h)\mathbb{E}[L_{t+\tau-1} | \mathbf{x}_t] + h(U-\mu)). \quad (32)$$

In numerical computations, we find that the required simulation run-length to obtain a given precision is around two orders of magnitude smaller than the required run-length when using the obvious estimator  $\frac{1}{t_2-t_1} \sum_{t=t_1+1}^{t_2} (pL_t + hJ_t)$ .

## 7. Numerical Results

The PIL policy is asymptotically optimal for large  $\tau$  if  $D$  has an exponential distribution. In Section 7.1, we provide numerical evidence that the PIL also outperforms the COP (and base-stock policy) for non-exponential distributions even for long lead times. Then in Section 7.2, we benchmark the performance of the PIL policy against other policies including the optimal policy for the standard



test-bed of Zipkin (2008a). Finally in Section 7.3, we provide numerical benchmarks for a test-bed of instances of the size one is likely to encounter in practice. All performance evaluations are done with simulation implemented in C. When multiple policies are compared, we use common random numbers. After a warm-up simulation run, run-length and batch-sizes in the simulation are set such that the half-width of a 95%-confidence interval is less than 1% of the point estimate, and there is no statistically significant dependence between consecutive batches. This section also provides results on the computational feasibility of computing inventory projections.

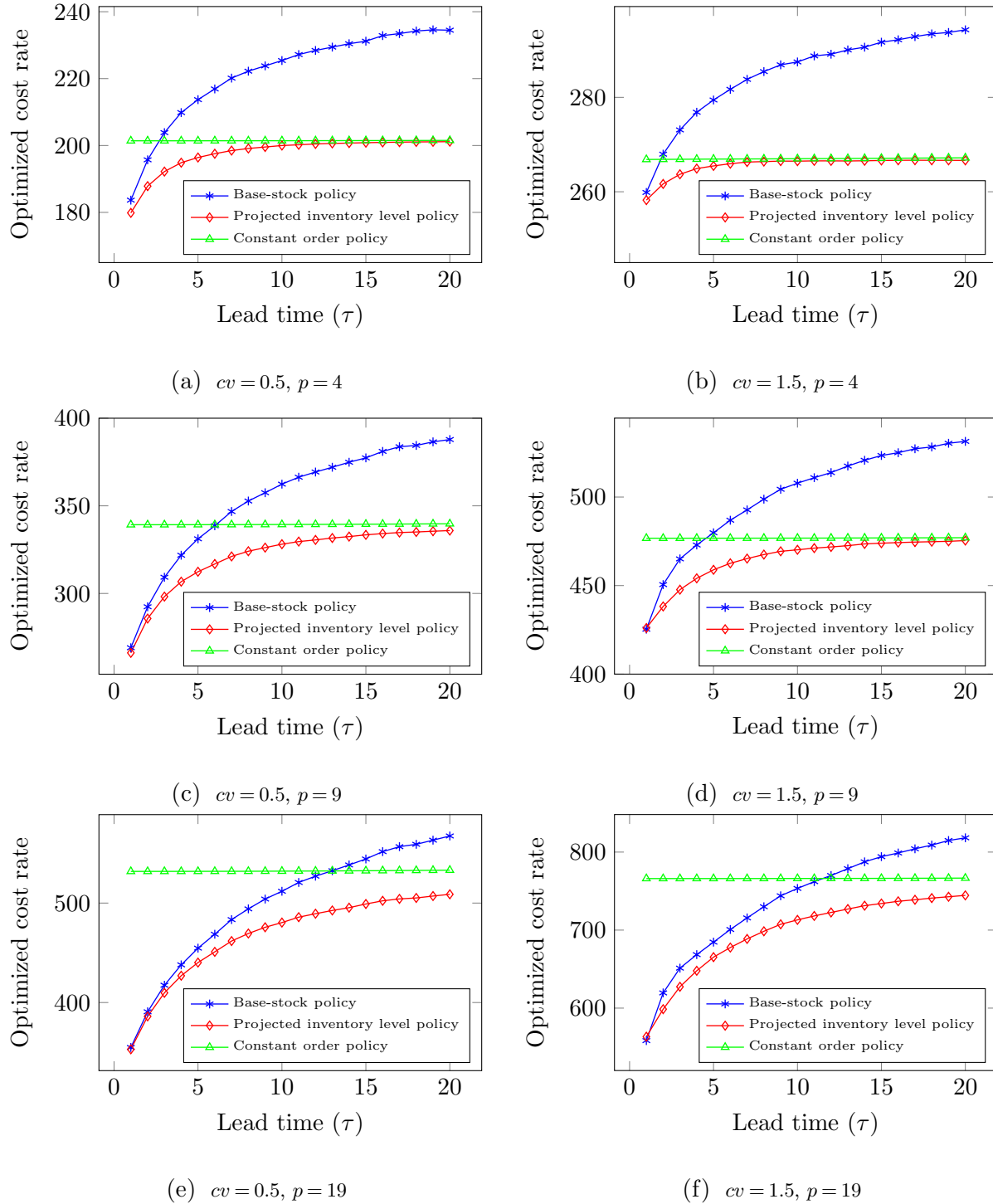
### 7.1. Long lead times

We created a test-bed in which demand has a Mixed Erlang distribution with mean 100 and the holding cost is fixed at  $h = 1$ . We then varied the coefficient of variation of the one period demand ( $\sqrt{\text{Var}[D]/(\mathbb{E}[D])^2}$ ) to be either  $\frac{1}{2}$  or  $\frac{3}{2}$  using the two moment fitting procedure in Section EC.7. We also varied the lead-time to be anything between 1 and 20,  $\tau \in \{1, \dots, 20\}$  and the penalty cost in the set  $p \in \{4, 9, 19\}$ . This makes for a total of  $2 \cdot 3 \cdot 20 = 120$  instances. For each of these instances we optimized the constant order policy, the base-stock policy and the PIL policy. The results are shown in Figure 1.

The PIL policy has superior performance relative to the COP and base-stock policy in all cases. Figure 1 also suggests that the PIL policy is asymptotically optimal as  $\tau \rightarrow \infty$  for demand distributions other than the exponential distribution.

### 7.2. Standard test-bed

Zipkin (2008a) provides a test-bed to compare the performance of notable policies for the canonical lost sales inventory model. This test-bed has relatively small instances as the performance of all policies including the optimal policy are evaluated numerically. This test-bed has two demand distributions, Poisson and geometric, both with mean 5. The holding cost is fixed at  $h = 1$  and the other parameters are varied as a full factorial:  $\tau \in \{1, 2, 3, 4\}$ ,  $p \in \{4, 9, 19, 39\}$  leading to a total of 32 instances. Zipkin (2008a) report the performance of notable policies advocated in literature



**Figure 1** Optimized Cost rate as a function of lead time for different policies for Mixed Erlang demand with mean 100 and coefficient of variation ( $cv$ ) of 0.5 (sub-figures a, c, and e) and 1.5 (sub-figures b, d, and f) for penalty costs of 4, 9 and 19 respectively and holding cost rate 1.

(e.g. Morton (1971), Levi et al. (2008), Huh et al. (2011)). Here we report on the base-stock policy, constant order policy, myopic policy, capped base-stock policy, and the PIL policy. The myopic policy is the best performing policy in Zipkin’s test-bed that has intuitive appeal. The myopic policy places an order in period  $t$  to minimize the projected cost in period  $t + \tau$  given the current state. The myopic policy is defined formally as:

$$M_t(\mathbf{x}) := \operatorname{argmin}_{q \geq 0} \mathbb{E}[pL_\tau + hJ_\tau \mid \mathbf{x}_0 = \mathbf{x}].$$

The capped base-stock policy is formally defined as

$$R_t^{S,r}(\mathbf{x}) := \min\{B_t^S(\mathbf{x}), C_t^r(\mathbf{x})\}.$$

The capped-base-stock policy is also asymptotically optimal both as  $p \rightarrow \infty$  and as  $\tau \rightarrow \infty$ , as is intuitively clear as the parameters of this policy can be set to mimic either a base-stock policy (by setting the cap  $r$  arbitrarily high) or as a constant order policy (by setting the base-stock level  $S$  arbitrarily high). Table 1 reports the performance of these policies. The performance of the best capped base-stock policies reported in Table 1 are taken from Xin (2020). He found the optimized parameters by using Matlab’s solver “fmincon”. We also replicated these results but found that whether we found the same policy performance depends on the initial solution provided to the solver. This indicates that the non-convexity of  $C(R^{S,r})$  in  $S$  and  $r$  can be a challenge.

The performance of the PIL policy is closest to optimal with an average optimality gap of 0.6% whereas the base-stock policy has an average optimality gap of 3.5%, the myopic policy of 2.8% and the capped base-stock policy of 0.7%. The average performance of the best constant order policy is quite poor with an average optimality gap of 47.4%. It appears that the PIL policy has attractive asymptotic properties as well as superior empirical performance compared to state of the art heuristics.

### 7.3. Large instance test-bed

We created a large test-bed of instances for which the optimal policy cannot be tractably computed. However, we believe these instances give a fair representation of instances that one may encounter in

**Table 1** Comparison of policies on Zipkin’s test-bed

Penalty per lost sale	Policy	Poisson demand				Geometric demand			
		Lead-time $\tau$				Lead-time $\tau$			
		1	2	3	4	1	2	3	4
$p = 4$	Optimal	4.04	4.40	4.60	4.73	9.82	10.24	10.47	10.61
	PIL	4.04	4.40	4.62	4.74	9.84	10.28	10.51	10.64
	Myopic	4.11	4.56	4.84	5.06	9.95	10.57	10.99	11.31
	Base-stock	4.16	4.64	4.98	5.20	10.04	10.70	11.13	11.44
	Capped base-stock	4.06	4.41	4.63	4.80	9.87	10.32	10.51	10.70
	COP	5.27				11.00			
$p = 9$	Optimal	5.44	6.09	6.53	6.84	14.51	15.50	16.14	16.58
	PIL	5.45	6.12	6.58	6.90	14.55	15.60	16.27	16.73
	Myopic	5.45	6.22	6.80	7.20	14.64	15.93	16.86	17.61
	Base-stock	5.55	6.32	6.86	7.27	14.73	15.99	16.87	17.54
	Capped base-stock	5.48	6.12	6.62	6.91	14.58	15.63	16.27	16.73
	COP	10.27				18.19			
$p = 19$	Optimal	6.68	7.66	8.36	8.89	19.22	20.89	22.06	22.95
	PIL	6.68	7.68	8.42	8.95	19.28	21.03	22.73	23.85
	Myopic	6.69	7.77	8.56	9.18	19.37	21.30	22.79	24.02
	Base-stock	6.73	7.84	8.60	9.23	19.40	21.31	22.73	23.85
	Capped base-stock	6.69	7.72	8.40	8.95	19.32	21.06	22.27	23.28
	COP	15.78				28.60			
$p = 39$	Optimal	7.84	9.11	10.04	10.79	23.87	26.21	27.96	29.36
	PIL	7.84	9.12	10.09	10.91	23.94	26.37	28.18	29.72
	Myopic	7.88	9.16	10.17	11.04	23.97	26.55	28.61	30.31
	Base-stock	7.86	9.19	10.22	11.06	24.00	26.55	28.51	30.12
	Capped base-stock	7.84	9.14	10.08	10.88	24.00	26.30	28.28	29.76
	COP	18.21				36.73			

practice. In all these instances, demand has a Mixed Erlang distribution with  $\mathbb{E}[D] = 100$  and  $h = 1$ . We varied the lead-time  $\tau \in \{1, 2, 3, 4, 5, 6\}$ , the penalty cost parameter  $p \in \{1, 4, 9, 19, 49, 99\}$ , and the coefficient of variation of the one period demand  $\sqrt{\text{Var}[D]/(\mathbb{E}[D])^2} \in \{0.15, 0.25, 0.5, 1, 1.5, 2.0\}$  for a total of 216 instances. (We use the two moment fitting procedure in Section EC.7 to fit a ME distribution.) We use the Matlab solver “fminbnd” to optimize the single parameter of the base-stock, and COP for each of these instances. Matlab’s multi-dimensional solver “fmincon” is used to optimize the parameters of the capped base-stock and the PIL policy with the following parameters that deviate from default: `FiniteDifferenceStepSize`= $10^{-2}$ , `OptimalityTolerance`= $10^{-3}$ , and `StepTolerance`= $10^{-8}$ . These changes to default settings have been made for the benefit of optimizing the capped base-stock policy as the estimator  $\widehat{C(R^{S,r})} = \frac{1}{t_2 - t_1} \sum_{t=t_1+1}^{t_2} (pL_t + hJ_t)$  is not smooth in  $S$  and  $r$  along a given sample path for the capped base-stock policy. Note that “fmincon” is the solver also used by Xin (2020) to optimize the parameters of the capped base-stock policy and represents the state of the art.  $C(R^{S,r})$  is not convex in  $S$  and  $r$ , and so there is no guarantee

that the best capped stock policy is found, whereas this is guaranteed for the PIL policy due to Theorem 5.

Both the PIL and the capped base-stock policy ran an initial optimization with a smaller simulation run-time ( $100\times$  smaller) to provide a good initial solution before running at high simulation precision. The PIL policy has the best performance of these four policies. Therefore we report the gap of the COP, base-stock and capped base-stock policies relative to the best PIL policy across this test-bed aggregated by setting; see Table 2. The performance of the PIL policy is again superior to the base-stock and constant order policy by a considerable margin on average (3.64% and 38.7%). The gap with the capped base-stock policy exceeds the base-stock policy at 3.65% on average. As the capped base-stock policy includes the base-stock policy as a special case, this indicates that “fmincon” cannot always find the best parameters due to the non-convexity of  $C(R^{S,r})$ . Running the optimization of the capped base-stock policy with multiple initial solutions improves the performance such that the average gap with the best PIL policy drops from 3.65% to 0.99%; details on this can be found in Table 2 of van Jaarsveld and Arts (2021). This approach more than doubles the computational time of optimizing the parameters of the capped base-stock policy.

The simulation of a given PIL policy requires the evaluation of the projected inventory level,  $\mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t]$ , for each period. Our implementation in C allowed us to do 18 million projections per minute on average for this test-bed. The computation times to optimize policy parameters for different policies are shown in Table 3. The average time to optimize a PIL policy is slightly less than the time to optimize the capped base-stock policy at 21.41 seconds and 22.91 seconds respectively. Optimization of the capped base-stock policy requires search in two dimensions and long simulation run-times for accuracy. The PIL policy requires projection, but only search in one dimension and shorter simulation run-times due to the low variance estimator in Section 6.3. Both policies have comparable computational requirements for the given test-bed. We do observe that the computation of optimal PIL-levels scales poorly in the lead-times whereas the computation of parameters for the capped base-stock policy remains unchanged as the lead time grows.

**Table 2** Comparison of policy performance on large test-bed

	Policy	Percentage gap with PIL policy								
		base-stock			constant order			capped base-stock		
	Gap with PIL	min	max	avg	min	max	avg	min	max	avg
CV of demand	0.4	0.08	16.49	4.77	0.10	199.53	49.20	0.10	16.49	4.79
	0.6	0.01	14.25	4.42	0.09	176.10	44.00	0.03	14.25	4.43
	0.8	-0.09	11.94	4.05	0.02	160.39	40.12	-0.09	11.95	4.05
	1	-0.15	10.17	3.49	-0.04	138.85	35.00	-0.14	10.17	3.49
	1.2	-3.81	8.61	2.83	0.02	125.56	31.70	-3.80	8.61	2.83
	1.4	-11.23	6.71	2.31	-0.02	123.77	32.01	-11.23	6.72	2.32
lead time ( $\tau$ )	1	-11.23	5.28	-0.18	0.52	199.53	57.97	-11.23	5.30	-0.17
	2	0.10	9.10	2.55	0.17	158.61	46.90	0.15	9.12	2.55
	3	0.24	11.80	3.60	0.07	133.20	38.72	0.26	11.80	3.61
	4	0.35	13.73	4.52	0.05	116.55	33.28	0.35	13.75	4.52
	5	0.42	15.23	5.33	-0.02	104.84	29.20	0.42	15.23	5.34
	6	0.83	16.49	6.05	-0.04	94.97	25.95	0.84	16.49	6.06
Penalty cost ( $p$ )	1	1.69	16.49	7.83	-0.04	4.10	0.55	1.69	16.49	7.84
	4	0.02	12.90	6.58	0.26	20.75	5.25	0.02	12.90	6.58
	9	-8.65	8.10	3.97	4.20	42.64	14.66	-8.65	8.10	3.98
	19	-11.23	4.72	2.02	14.63	74.15	31.79	-11.23	4.72	2.03
	49	-8.79	4.82	0.88	40.57	133.93	69.04	-8.79	4.82	0.89
	99	-7.19	4.93	0.58	71.85	199.53	110.73	-7.18	4.93	0.59
Total		-11.23	16.49	3.64	0.0	199.5	38.7	-11.23	16.49	3.65

**Table 3** Computation times to optimize policy parameters

		Computation times in seconds											
		base-stock			constant order			capped base-stock			PIL		
		min	max	avg	min	max	avg	min	max	avg	min	max	avg
CV of demand	0.4	7.49	12.60	9.98	8.05	12.20	10.37	3.10	56.33	22.35	1.43	118.15	34.36
	0.6	6.77	11.76	9.03	6.67	10.73	8.93	2.88	56.42	20.49	1.10	69.10	21.86
	0.8	6.33	12.34	9.44	6.94	10.90	8.85	13.37	99.20	25.89	0.56	97.88	16.08
	1	7.04	11.62	8.93	6.68	10.33	8.27	3.08	51.90	23.48	0.50	40.19	11.02
	1.2	7.63	11.86	9.19	6.04	9.83	7.71	2.90	58.36	22.68	0.59	56.85	15.32
	1.4	5.89	11.16	9.27	5.15	9.88	7.18	2.97	65.91	22.58	0.63	150.78	29.83
lead time ( $\tau$ )	1	5.89	10.01	8.26	5.19	12.19	8.57	2.90	39.22	20.81	0.56	10.55	2.43
	2	7.69	11.04	9.18	5.18	12.15	8.59	2.90	56.42	25.49	0.50	25.08	6.39
	3	7.69	11.86	9.30	5.15	12.15	8.62	3.13	99.20	24.35	1.06	30.00	14.25
	4	7.52	11.76	9.55	5.21	12.13	8.54	2.97	65.91	23.09	1.56	74.15	22.13
	5	6.82	12.03	9.64	5.16	11.19	8.42	3.08	41.43	22.27	2.42	125.91	34.79
	6	7.67	12.60	9.91	5.16	12.20	8.56	2.88	40.14	21.46	3.53	150.78	48.49
Penalty cost ( $p$ )	1	6.33	11.18	8.86	5.15	8.16	6.65	11.69	43.13	23.75	0.50	11.59	3.23
	4	6.77	10.21	8.51	5.76	10.00	7.77	2.90	56.42	22.00	1.20	88.42	14.70
	9	5.89	9.63	8.47	6.51	11.47	8.46	2.90	58.36	25.44	0.56	101.76	22.33
	19	7.36	10.80	9.27	7.29	11.43	8.84	3.09	99.20	21.18	0.75	125.91	27.35
	49	8.39	11.76	10.32	8.14	12.19	9.64	12.82	65.91	23.54	1.98	150.78	29.47
	99	6.82	12.60	10.41	5.94	12.20	9.95	2.88	45.53	21.56	0.98	118.15	31.40
Total		5.89	12.60	9.31	5.1	12.2	8.6	2.88	99.20	22.91	0.50	150.78	21.41

## 7.4. Discussion

A key finding of Sections 7.1-7.3 is that the PIL policy consistently outperforms the COP also for non-exponential demand distributions, while the corresponding rigorous proof of Section 4.1

crucially assumes exponentially distributed demand. We briefly hypothesize in an attempt to understand and partially reconcile the difference in scope.

Recall that we established dominance of the PIL policy over the COP policy for exponential demand, by showing that the PIL policy can be obtained by performing a policy improvement step on the COP policy. For general demand, the 1-step policy improvement  $\pi^{r,+}$  on the constant order policy must satisfy  $\pi^{r,+}(\mathbf{x}_t) \in \operatorname{argmin}_{q \geq 0} \mathbb{E}[\mathcal{H}^r(J_{t+\tau-1} + q) \mid \mathbf{x}_t]$ , where  $\mathcal{H}^r(\cdot)$  represents the bias associated with  $C^r$ . For exponentially distributed demand, the bias exists and takes the form of a parabola: As a consequence,  $\pi^{r,+}(\mathbf{x}_t)$  depends only on the expectation of  $J_{t+\tau-1}$  and takes the simple form of the PIL policy. However,  $\mathcal{H}^r(\cdot)$  will be some general convex function for other demand distributions (provided existence). As a consequence  $\pi^{r,+}$  will depend on the precise distribution of  $J_{t+\tau-1}$ , i.e. it will not be a PIL policy.

However, in line with common approaches in approximate dynamic programming (Powell 2007), one might consider to approximate the bias, and in view of the convexity of  $\mathcal{H}^r(\cdot)$ , it would be reasonable to fit a second-degree polynomial. Such an approximation would yield an *approximate*  $\pi^{r,+}$  that is a PIL policy; arguably, our numerical results indicate that such approximations tend to work well.

## 8. Conclusion

We introduced the projected inventory level policy for the lost sales inventory system. We showed that this policy is asymptotically optimal in two regimes and has superior numerical performance. The policy may also be applied to the back-order system, in which case it is equivalent to a base-stock policy.

It needs to be explored whether projected inventory level policies can be fruitfully used for other complicated inventory systems where the whole pipeline is relevant for ordering decisions. Such systems include perishable inventory systems (c.f. Bu et al. 2022), dual sourcing systems (e.g. Xin and Goldberg 2018), and systems with independent (overtaking) lead-times (e.g. Stolyar and Wang 2021).

## Acknowledgments

The authors are grateful to the department editor, the associate editor, and all the anonymous referees for a thoughtful and constructive review process. The authors thank an anonymous referee for providing the key idea underlying the proof of Lemma 9, and Ton de Kok, Ivo Adan and Geert-Jan van Houtum for stimulating discussions.

## References

- Agrawal, Shipra, Randy Jia. 2022. Learning in structured MDPs with convex cost functions: Improved regret bounds for inventory management. *Operations Research* **70**(3) 1646–1664.
- Bai, Xingyu, Xin Chen, Menglong Li, Alexander Stolyar. 2023. Asymptotic optimality of semi-open-loop policies in markov decision processes with large lead times. *Operations Research* **Ahead of Print**.
- Bijvank, M., W.T. Huh, G. Janakiraman, W. Kang. 2014. Robustness of order-up-to policies in lost-sales inventory systems. *Operations Research* **62**(5) 1040–1047.
- 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**(1) 106–114.
- Bijvank, M., I.F.A. Vis. 2011. Lost-sales inventory theory: A review. *European Journal of Operational Research* **215**(1) 1–13.
- Bu, Jinzhi, Xiting Gong, Xiuli Chao. 2022. Asymptotic optimality of base-stock policies for perishable inventory systems. *Management Science* **articles in advance**.
- Bu, Jinzhi, Xiting Gong, Dacheng Yao. 2020. Constant-order policies for lost-sales inventory models with random supply functions: Asymptotics and heuristic. *Operations Research* **68**(4) 1063–1073.
- Chen, Wei, Milind Dawande, Ganesh Janakiraman. 2014. Fixed-dimensional stochastic dynamic programs: An approximation scheme and an inventory application. *Operations Research* **62**(1) 81–103.
- Esary, J.D., F. Prochan, D.W. Walkup. 1967. Association of random variables with applications. *The Annals of Mathematical Statistics* **38**(5) 1466–1474.
- 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.



- Goldberg, D.A., M.I. Reiman, Q. Wang. 2021. A survey of recent progress in the asymptotic analysis of inventory systems. *Production and Operations Management* **30**(6) 1718–1750.
- Haijema, René, Jan van der Wal, et al. 2008. An mdp decomposition approach for traffic control at isolated signalized intersections. *Probability in the Engineering and Informational Sciences* **22**(4) 587–602.
- Huh, Woonghee Tim, Ganesh Janakiraman, Mahesh Nagarajan. 2011. Average cost single-stage inventory models: An analysis using a vanishing discount approach. *Operations Research* **59**(1) 143–155.
- Huh, W.T., G. Janakiraman, J.A. Muckstadt, P. Rusmevichientong. 2009a. An adaptive algorithm for finding the optimal base-stock policy in lost sales inventory systems with censored demand. *Mathematics of Operations Research* **34**(2) 397–416.
- Huh, W.T., G. Janakiraman, J.A. Muckstadt, P. Rusmevichientong. 2009b. Asymptotic optimality of order-up-to policies in lost sales inventory systems. *Management Science* **55**(3) 404–420.
- Janakiraman, G., S. Seshadri, G.J. Shanthikumar. 2007. A comparison of the optimal costs of two canonical inventory systems. *Operations Research* **55**(5) 866–875.
- Johansen, Søren Glud, Anders Thorstenson. 2008. Pure and restricted base-stock policies for the lost-sales inventory system with periodic review and constant lead times. *15th International Symposium on Inventories*.
- Karlin, S., H. Scarf. 1958. Inventory models of the Arrow-Harris-Marchak 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.
- Levi, R., G. Janakiraman, M. Nagaraja. 2008. A 2-approximation algorithm for stochastic inventory models with lost sales. *Mathematics of Operations Research* **33**(2) 351–374.
- Morton, K. 1969. Bounds on the solution of the lagged optimal inventory equation with no demand backlogging and proportional costs. *SIAM Review* **11**(4) 572–596.
- Morton, K. 1971. The near-myopic nature of the lagged-proportional-cost inventory problem with lost sales. *Operations Research* **19**(7) 1708–1716.
- Powell, Warren B. 2007. *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703. John Wiley & Sons.

- Song, Jing-Sheng. 1998. On the order fill rate in a multi-item, base-stock inventory system. *Operations Research* **46**(6) 831–845.
- Stolyar, A.L., Q. Wang. 2021. Exploiting random lead times for significant inventory cost savings. *Operations Research* **70**(4) 2496–2516.
- Sun, Peng, Kai Wang, Paul Zipkin. 2014. Quadratic approximation of cost functions in lost sales and perishable inventory control problems. *Fuqua School of Business, Duke University, Durham, NC* .
- Tijms, H.C. 2003. *A First Course in Stochastic Models*. John Wiley & Sons.
- van Donselaar, K., T. de Kok, W. Rutten. 1996. Two replenishment strategies for the lost sales inventory model: A comparison. *International Journal of Production Economics* **46-47** 285–295.
- van Houtum, G.J. 2006. Multiechelon production/inventory systems: optimal policies, heuristics and algorithms. *Tutorials in Operations Research INFORMS* 163–199.
- van Jaarsveld, W. 2020. Deep controlled learning of dynamic policies with an application to lost-sales inventory control. *arXiv preprint arXiv:2011.15122* .
- van Jaarsveld, W., J. Arts. 2021. Projected inventory level policies for lost sales inventory systems: Asymptotic optimality in two regimes. *arXiv:2101.07519v2* .
- Xin, Linwei. 2020. Understanding the performance of capped base-stock policies in lost-sales inventory models. *Operations Research* **69**(1) 61–70.
- Xin, Linwei, David 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, Linwei, David A Goldberg. 2018. Asymptotic optimality of tailored base-surge policies in dual-sourcing inventory systems. *Management Science* **64**(1) 437–452.
- Zhang, Huanan, Xiuli Chao, Cong Shi. 2020. Closing the gap: A learning algorithm for lost-sales inventory systems with lead times. *Management Science* **66**(5) 1962–1980.
- Zipkin, P. 2008a. Old and new methods for lost-sales inventory systems. *Operations Research* **56**(5) 1256–1263.
- Zipkin, P. 2008b. On the structure of lost-sales inventory models. *Operations Research* **56**(4) 937–944.

## Technical proofs for companion

### EC.1. Proof of Lemma 1

*Proof.* Suppose actions follow a PIL policy  $P_t^U(\mathbf{x}_t)$  for given  $U \geq 0$ . Suppose the state  $\mathbf{x}_0$  of period 0 satisfies  $\mathbb{E}[J_{\tau-1}|\mathbf{x}_0] \leq U$ . Denote by  $q_\tau \geq 0$  the non-negative order placed in period 0 that attains  $\mathbb{E}[I_\tau|\mathbf{x}_0] = \mathbb{E}[J_{\tau-1} + q_\tau|\mathbf{x}_0] = U$ .

Conditional on  $D_0$  being  $d_0$ , the projected inventory level as seen in period 1 before the order  $q_{\tau+1}$  is placed is given by  $\mathbb{E}[J_\tau | D_0 = d_0]$ . It is sufficient to show that

$$\forall d_0 \geq 0 : \mathbb{E}[J_\tau | \mathbf{x}_0, D_0 = d_0] \leq U, \quad (\text{EC.1})$$

since this implies that for period 1 it is always possible to place a non-negative order  $q_{\tau+1}$  to attain the projected inventory level  $U$ . (Periods 2, 3, ... then follow by induction.) Since  $J_\tau$  is decreasing in  $D_0$  for all  $D_1, \dots, D_\tau$  we find  $\mathbb{E}[J_\tau | D_0 = d_0] \leq \mathbb{E}[J_\tau | D_0 = 0]$ , and hence to show (EC.1) it suffices to demonstrate that  $\mathbb{E}[J_\tau | D_0 = 0] \leq \mathbb{E}[I_\tau] = U$ , which will be the subject of the remainder of this proof.

Define the random variable  $G(y)$  such that  $\mathbb{P}(G(y) \leq g) = \mathbb{P}(I_\tau \leq g | \mathbf{x}_t, J_0 = y)$ , or equivalently  $G(y) = ((\dots((y + q_1 - D_1)^+ + q_2 - D_2)^+ \dots)^+ + q_{\tau-1} - D_{\tau-1})^+ + q_\tau$ . Observe that we have  $\mathbb{E}[I_\tau|\mathbf{x}_0] = \mathbb{E}[G((I_0 - D)^+)]$  and  $\mathbb{E}[J_\tau | D_0 = 0] = \mathbb{E}[(G(I_0) - D)^+]$ . Note that  $0 \leq dG(y)/dy \leq 1$  and  $G(y) \geq 0$  so that

$$(G(I_0) - D)^+ \leq G((I_0 - D)^+). \quad (\text{EC.2})$$

It follows from (EC.2) that  $\mathbb{E}[J_\tau | D_0 = 0] = \mathbb{E}[(G(I_0) - D)^+] \leq \mathbb{E}[G((I_0 - D)^+)] = \mathbb{E}[I_\tau]$ .  $\square$

### EC.2. Proof of Lemma 2

*Proof.* First note that the inventory at the end of a period under  $C^r$  satisfies the Lindley recursion  $J_{t+1} = (J_t + r - D_t)^+$  so that  $\lim_{t \rightarrow \infty} \mathbb{E}[J_t]$  can be determined with the Pollaczek Khinchine equation for the mean waiting time in an  $M/D/1$  queue. Therefore we can express  $g^r$  in closed form as

$$g^r = p(\mu - r) + h \frac{r^2}{2(\mu - r)}. \quad (\text{EC.3})$$

We can now directly verify that inserting (7) and (EC.3) into the right hand side of (6) again equals (7) by using integration by parts and tedious but otherwise straightforward algebra:

$$\begin{aligned}
& \mathbb{E}_D [h(x-D)^+ + p(D-x)^+ + \mathcal{H}^r((x-D)^+ + r)] - g^r \\
&= p\mathbb{E}[(D-x)^+] + h\mathbb{E}[(x-D)^+] + \int_0^x \frac{\mathcal{H}^r(x-y+r)}{\mu} e^{-y/\mu} dy + \mathcal{H}^r(r)e^{-x/\mu} - g^r \\
&= p\mu e^{-x/\mu} + h(x-\mu + \mu e^{-x/\mu}) + e^{-x/\mu} x \left( \frac{h(3r^2 + 3rx + x^2)}{6\mu(\mu-r)} + \frac{pr(2r+x)}{2\mu(\mu-r)} - \frac{p/(2r+x)}{2(\mu-r)} \right) \\
&\quad + e^{-x/\mu} \left( \frac{hr^2}{2(\mu-r)} - pr^2 \right) - p(\mu-r) - h \frac{r^2}{2(\mu-r)} \\
&= \frac{h}{2(\mu-r)} x^2 - px = \mathcal{H}^r(x). \tag{EC.5}
\end{aligned}$$

Clearly  $\mathcal{H}^r(x) = \frac{h}{2(\mu-r)} x^2 - px$  also satisfies  $\mathcal{H}^r(0) = 0$ .  $\square$

### EC.3. Proof of Lemma 4

*Proof.* We prove the following statement by induction on  $t_2$ , starting at  $t_2 = t_1$ :

$$\mathcal{H}^r(I_{t_1}) = \mathbb{E}_{D_{t_1}, \dots, D_{t_2}} [c[t_1, t_2](C^r) + \mathcal{H}^r(I_{t_2+1}) | I_{t_1}] - (t_2 + 1 - t_1)g^r. \tag{EC.6}$$

The base case  $t_2 = t_1$  holds by (6), the definition of  $c_t$ , and  $I_{t_1+1} = (I_{t_1} - D_{t_1}) + r$ . Now for the inductive step, assume that the statement holds for some  $t_2 \geq t_1$ ; we will show it holds also for  $t_2 + 1$ . Use (6) to conclude that

$$\begin{aligned}
\mathbb{E}_{D_{t_1}, \dots, D_{t_2}} [\mathcal{H}^r(I_{t_2+1}) | I_{t_1}] &= \mathbb{E}_{D_{t_1}, \dots, D_{t_2}} \left[ \mathbb{E}_{D_{t_2+1}} [c_{t_2+1} + \mathcal{H}^r((I_{t_2+1} - D_{t_2+1})^+ + r) - g^r | I_{t_2+1}] | I_{t_1} \right] \\
&= \mathbb{E}_{D_{t_1}, \dots, D_{t_2}} \left[ \mathbb{E}_{D_{t_2+1}} [c_{t_2+1} + \mathcal{H}^r(I_{t_2+2}) | I_{t_2+1}] | I_{t_1} \right] - g^r \\
&= \mathbb{E}_{D_{t_1}, \dots, D_{t_2}, D_{t_2+1}} [c_{t_2+1} + \mathcal{H}^r(I_{t_2+2}) | I_{t_1}] - g^r. \tag{EC.7}
\end{aligned}$$

Now substitution of (EC.7) back into the induction hypothesis (EC.6) yields:

$$\begin{aligned}
\mathcal{H}^r(I_{t_1}) &= \mathbb{E}_{D_{t_1}, \dots, D_{t_2}} [c[t_1, t_2](C^r) + \mathcal{H}^r(I_{t_2+1}) | I_{t_1}] - (t_2 + 1 - t_1)g^r \\
&= \mathbb{E}_{D_{t_1}, \dots, D_{t_2+1}} [c[t_1, t_2 + 1](C^r) + \mathcal{H}^r(I_{t_2+2}) | I_{t_1}] - (t_2 + 2 - t_1)g^r. \tag{EC.8}
\end{aligned}$$

By induction, this shows that (EC.6) holds for all  $t_2 \geq t_1$ .  $\square$

## EC.4. Proof of Lemma 5

*Proof.* We have

$$\mathbb{E}[(X + Y)^+] = \int_{x+y \geq 0} (x + y) dF(x, y) = \int_{x=-\infty}^{\infty} \int_{y=-x}^{\infty} (x + y) dF(x, y). \quad (\text{EC.9})$$

For  $x + y \geq 0$  we can write

$$x + y = \int_{z=-y}^x dz. \quad (\text{EC.10})$$

Substitution of (EC.10) into (EC.9) yields

$$\begin{aligned} \mathbb{E}[(X + Y)^+] &= \int_{x=-\infty}^{\infty} \int_{y=-x}^{\infty} \int_{z=-y}^x dz dF(x, y) = \int_{-y \leq z \leq x} dF(x, y) dz \\ &= \int_{z=-\infty}^{\infty} \int_{x=z}^{\infty} \int_{y=-z}^{\infty} dF(x, y) dz = \int_{z=-\infty}^{\infty} \mathbb{P}(X \geq z, Y \geq -z) dz. \quad \square \end{aligned}$$

## EC.5. Proof of Lemma 9

*Proof.* We first derive an identity for  $q[1, T + \tau](P^U) | D_0, \dots, D_{T-1}$ . From (1-3), we find  $J_t - L_t = I_t - D_t = J_{t-1} + q_t - D_t$ , which implies  $J_t - J_{t-1} = L_t + q_t - D_t$ . Summing this identity for  $t = 1, \dots, T + \tau - 1$  yields:  $\sum_{t=1}^{T+\tau-1} J_t - J_{t-1} = J_{T+\tau-1} - J_0 = L[1, T + \tau - 1] + q[1, T + \tau - 1] - D[1, T + \tau - 1]$ . Since  $J_0 = L_0 + I_0 - D_0$  by (1-2), we obtain

$$J_{T+\tau-1} = I_0 + L[0, T + \tau - 1] + q[1, T + \tau - 1] - D[0, T + \tau - 1] \quad (\text{EC.11})$$

Next, note that  $\mathbb{E}(J_{T+\tau-1}(P^U) | D_0, \dots, D_{T-1}) = \mathbb{E}(J_{T+\tau-1}(P^U) | \mathbf{x}_T)$ , since  $\mathbf{x}_T$  is known given  $D_t, \dots, D_{T-1}$ . Thus, Lemma 1 and (4) imply  $\mathbb{E}(J_{T+\tau-1}(P^U) | D_0, \dots, D_{T-1}) = U - q_{T+\tau}(P^U) | D_0, \dots, D_{T-1}$ . Taking the conditional expectation on both sides of (EC.11) and plugging in this latter identity yields  $U - q_{T+\tau}(P^U) | D_0, \dots, D_{T-1} = I_0 + \mathbb{E}(L[0, T + \tau - 1] + q[1, T + \tau - 1] - D[0, T + \tau - 1] | D_0, \dots, D_{T-1})$ . Using  $\mathbb{E}(D[0, T + \tau - 1] | D_0, \dots, D_{T-1}) = D[0, T - 1] + \tau\mu$  and rearranging terms yields:

$$\begin{aligned} &q[1, T + \tau](P^U) | D_0, \dots, D_{T-1} \\ &= U - I_0 - D[0, T - 1] - \tau\mu - \mathbb{E}[L[1, T + \tau - 1](P^U) | D_0, \dots, D_{T-1}] \end{aligned} \quad (\text{EC.12})$$

We are now ready to prove Properties 1 and 2 by induction. As induction hypothesis, let  $T \in \mathbb{N}$  and suppose that  $\forall t < T$  and for all demand sequences  $D_0, D_1, \dots$  the random variables  $q[1, t + \tau](P^U)|D_0, \dots, D_{t-1}$  are non-decreasing and concave, while the random variables  $L[1, t + \tau](P^U)|D_0, \dots, D_{t+\tau}$  are non-increasing and convex. We will prove that these random variables are also non-decreasing and concave (resp. non-increasing and convex) for  $t = T$ .

Since  $L[1, T + \tau - 1](P^U)|D_0, \dots, D_{T+\tau-1}$  is non-increasing and convex in  $U$  by induction hypothesis, and since these properties are preserved in expectation, we find that  $\mathbb{E}[L[1, T + \tau - 1](P^U)|D_0, \dots, D_{T+\tau-1}] = \mathbb{E}[L[1, T + \tau - 1](P^U)|D_0, \dots, D_{T-1}]$  is non-increasing and convex in  $U$  for every realization  $D_0, \dots, D_{T-1}$ . By (EC.12) and since  $U$  is non-decreasing and concave in  $U$ , this implies that  $q[1, T + \tau](P^U)|D_0, \dots, D_{T-1}$  is non-decreasing and concave in  $U$ . Next, specializing (17) to  $P^U$ , we find:

$$L[0, T + \tau](P^U) = \max_{k \in \{0, \dots, T + \tau\}} (D[0, k] - I_0 - q[1, k](P^U))^+, \quad (\text{EC.13})$$

Since  $L[0, T + \tau](P^U)$  is the maximum of a number of functions, that by induction hypothesis and by the result we just proved are decreasing and convex in  $U$ ,  $L[0, T + \tau](P^U)$  must also be decreasing and convex in  $U$ . This proves the inductive step.

Now note that our induction hypothesis holds for  $T = 1$  since in that case the functions are independent of  $U$ , and hence non-increasing and convex (non-decreasing and concave). This completes our proof by induction of Properties 1 and 2. For Property 3, note that it follows directly from Property 2.  $\square$

## EC.6. Explicit recursions for inventory projection

All probabilities below will be conditional on the state at time 0,  $\mathbf{x}_0$ . We omit this condition as it makes the derivations more readable. From equations (1)-(3), we obtain the following recursive expressions for any  $t \in \{0, \dots, \tau\}$ :

$$\mathbb{P}(\tilde{J}_t = x) = \sum_{k=0}^{\infty} \mathbb{P}((\tilde{I}_t - K_t)^+ = x | K_t = k) \mathbb{P}(K_t = k) = \sum_{k=0}^x \mathbb{P}(\tilde{I}_t = x - k) \theta_k \quad (\text{EC.14})$$

$$\mathbb{P}(\tilde{L}_t = x) = \sum_{k=0}^{\infty} \mathbb{P}((K_t - \tilde{I}_t)^+ = x \mid K_t = k) \mathbb{P}(K_t = k) = \sum_{k=x+1}^{\infty} \mathbb{P}(\tilde{I}_t = k - x) \theta_k \quad (\text{EC.15})$$

$$\mathbb{P}(\tilde{I}_t = x) = \mathbb{P}(\tilde{J}_{t-1} + Q_t = x) = \sum_{y=0}^{\infty} \mathbb{P}(\tilde{J}_{t-1} = x - y) \mathbb{P}(Q_t = y), \quad (\text{EC.16})$$

These recursions can be applied running from  $t = 0$  till  $t = \tau$  as  $\tilde{I}_0$  and  $Q_t$ ,  $t \in \{1, \dots, \tau - 1\}$  have Poisson distributions with means  $\lambda I_0$  and  $\lambda q_t$  conditional on  $\mathbf{x}_0 = (I_0, q_1, \dots, q_{\tau-1})$  when demand has a ME distribution with scale  $\lambda$ :

$$\mathbb{P}(\tilde{I}_0 = x) = e^{-\lambda I_0} \frac{(\lambda I_0)^x}{x!}, \quad \mathbb{P}(Q_t = x) = e^{-\lambda q_t} \frac{(\lambda q_t)^x}{x!}. \quad (\text{EC.17})$$

## EC.7. Two moment fits for mixed Erlang distributions

This Section provides the two moment fits that can also be found in Tijms (2003) or van Houtum (2006). Suppose that  $D := \sum_{i=1}^K E_i$  where  $\{E_i\}_{i=1}^{\infty}$  is a sequence of i.i.d. exponential random variables with rate  $\lambda$  and that  $K$  is an independent discrete random variable on the integers with mass function  $\theta_k = \mathbb{P}(K = k)$ . When the mean  $\mu = \mathbb{E}[D]$  of demand and the squared coefficient of variation  $c_D^2 = \text{Var}[D]/\mu^2$  are known, the following two point distribution for  $K_t$  and choice for  $\lambda$  will match these moments when  $c_D^2 \leq 1$ :

$$\theta_k = \begin{cases} p & \text{if } k = k_1 \\ 1 - p & \text{if } k = k_2 \\ 0 & \text{otherwise} \end{cases}, \quad \lambda = \frac{k_1 - p}{\mu},$$

with

$$k_1 := \left\lfloor \frac{1}{c_D^2} \right\rfloor, \quad k_2 = k_1 + 1, \quad p = \frac{k_1 c_D^2 - \sqrt{k_1(1 + c_D^2) - k_1^2 c_D^2}}{1 + c_D^2}.$$

When  $c_D^2 > 1$ , then an appropriate parameterization of a mixed Erlang distribution is

$$\theta_k = \begin{cases} p & \text{if } k = k_1 \\ 1 - p & \text{if } k = k_2 \\ 0 & \text{otherwise} \end{cases}, \quad \lambda = \frac{p + k_2(1 - p)}{\mu},$$

with

$$k_1 = 1, \quad k_2 = \max \left( 3, \left\lceil \frac{4c_D^2 + \sqrt{16(c_D^2 - 1)}}{2} \right\rceil \right), \quad p = \frac{k_1 c_D^2 - \sqrt{k_1(1 + c_D^2) - k_1^2 c_D^2}}{1 + c_D^2}.$$