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, 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 τ. 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 6), 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 (2019). 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 is 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.

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 where 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 2019). 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. (2019) 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. 2020b, Xin 2019, Bai et al. 2020). A natural question is whether any intuitive policies exist that are asymptotically optimal in both regimes. Xin (2019) propose the capped base-stock policy which was first introduces 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 (2019) 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:

$$L_t = (D_t - I_t)^+ \tag{1}$$

$$J_t = (I_t - D_t)^+ \tag{2}$$

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

where $(x)^+ := \max(0, x)$. The state of this system in period t is given by $\mathbf{x}_t = (I_t, q_{t+1}, q_{t+2}, \ldots, q_{t+\tau-1}) \in \mathbb{R}_+^\tau$. 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 $\mathrm{A} = \{\mathrm{A}_t, t \in \mathbb{N}_0\}$, 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:

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

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

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

$$\mathrm{P}_t^U(\mathbf{x}_t) := (U - \mathbb{E}[J_{t+\tau-1}|\mathbf{x}_t])^+ = \left(U - I_t - q[t+1, t+\tau-1] - \mathbb{E}\left[L[t, t+\tau-1]\big|\mathbf{x}\right] + \tau\mu\right)^+. \qquad (4)$$

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. *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 ($\mathbb{E}[J_{\tau-1}] \leq U$) for any given PIL policy.*

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 $c[a, b](\mathrm{A})$ to make explicit the dependence on A. The cost-rate associated with a policy A is then

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

Let $S^* \in \mathrm{argmin}_S\, C(\mathrm{B}^S)$, $r^* \in \mathrm{argmin}_r\, C(\mathrm{C}^r)$, and $U^* \in \mathrm{argmin}_U\, C(\mathrm{P}^U)$ 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 $f(q|\mathbf{x}_t) = \lim_{T\to\infty} \mathbb{E}[F^T(q) - F^T(r) \mid \mathbf{x}_t]$. 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 $\mathrm{P}^U(\mathbf{x}_t) \in \operatorname{argmin}_{q\geq 0} f(q|\mathbf{x}_t)$ 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 ϵ has two effects on the infinite horizon costs: 1) ϵ more demand is eventually satisfied (not lost) and a penalty cost $p\epsilon$ is averted 2) From time $t+\tau$ until the first stock-out, the inventory level increases by ϵ (for ϵ *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

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

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 $g^r = C(\mathrm{C}^r)$ be the long run average costs of the C^r policy. Then the bias $\mathcal{H}^r(\cdot) : \mathbb{R}^+ \to \mathbb{R}$ associated with C^r satisfies:

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

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.

Proposition 1. *When demand has an exponential distribution, the bias of the constant order policy with constant order quantity $r < \mu$ is a parabola:*

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

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

The proof of Proposition 1 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 2. *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 $q_{t+\tau} = \mathrm{P}^{U(r)}(\mathbf{x}_t)$ 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. For any $\mathbf{x}_t$ and associated $q_{t+\tau} = \mathrm{P}_t^{U(r)}(\mathbf{x}_t)$, we have $q_{t+\tau} = (U(r) - \mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t])^+$. To see this note that $J_{t+\tau-1} = I_t + q[t+1, t+\tau-1] + L[t, t+\tau-1] - D[t, t+\tau-1]$. Thus

$$\mathbb{E}[J_{t+\tau-1} \mid \mathbf{x}_t] = I_t + q[t+1, t+\tau-1] - \mathbb{E}\left[D[t, t+\tau-1] \mid \mathbf{x}_t\right] - \tau\mu$$

so that the claimed result follows from (4).

Using Lemma 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 $a_2 = -\left(\frac{p(\mu - r^*)}{h}\right)^2$, thus $U(r)$ is the unique minimizer of $\mathcal{H}^r(\cdot)$. Now observe that

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

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

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 (2) that the PIL dominates the constant order policy. Our proof uses the following result, the proof of which appears in Appendix EC.3:

Lemma 3. *Let $t_1 \leq t_2$, $t_1, t_2 \in \mathbb{N}_0$, $r \in [0, \mathbb{E}[D])$, $g^r = C(\mathrm{C}^r)$, and suppose $q_t = r$ for all $t \in \{t_1 + 1, \ldots, t_2\}$. Then*

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

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 $C(\mathrm{P}^{U^*}) \leq C(\mathrm{P}^{U(r^*)}) \leq C(\mathrm{C}^{r^*})$ for any $\tau \in \mathbb{N}_0$, where $U(r)$ is given by Lemma 2.*

Proof. We focus on bounding $\mathbb{E}[c[\tau,T](\mathrm{P}^{U(r^*)}) - c[\tau,T](C^{r^*})]$. 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 $\mathrm{P}(U(r^*))$, and subsequent orders using the optimal constant-order policy C^{r^*}:

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

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 $\mathcal{G}^{\tilde{t}}$:

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

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 inequality, note that $\mathcal{G}^{\tilde{t}+1}_t = \mathcal{G}^{\tilde{t}}_t = \mathrm{C}^{r^*}$ 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 3.

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

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

For the third equality, we use Lemma 2. For the fourth equality, observe that $\mathbf{x}_0 = 0$, and hence $\mathbb{E}[J_{\tilde{t}-1}|\mathbf{x}_{\tilde{t}}] \leq U(r^*)$ (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:

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

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

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

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 $\lim_{T\to\infty} \frac{1}{T-\tau+1}(\mathbb{E}[I_{T+1}(\mathrm{C}^{r^*})] - \mathbb{E}[I_{T+1}(\mathrm{P}^{U(r^*)})]) = 0$. This latter observation 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 $(\tau+1)U(r^*)$.

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

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

This completes the proof. □

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 of between the order decisions of both policies; see (12). The amount 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 $\tau \to \infty$

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

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

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

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 τ.

5. Penalty cost asymptotics

The performance of the PIL-policy in the asymptotic regime that $p \to \infty$ 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 backorder 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 backorder 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 backorder system

The backorder system with lead-time $\tau \in \mathbb{N}_0$ holding cost parameter $h > 0$ per period per item and backorder 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 backorder system are:

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

where B_t and $I_t^{\mathcal{B}}$ denote respectively the number of items on backorder 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:

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

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

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

The optimal average cost-rate for this system satisfies:

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

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 $c^{\mathcal{B}}[a,b](\mathrm{A})$ 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 $C^{\mathcal{B}}(\mathrm{A}) = \limsup_{T\to\infty} \mathbb{E}\left[\frac{1}{T-\tau+1} c^{\mathcal{B}}[\tau, T](\mathrm{A})\right]$. 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

$C^* \le C^{\mathcal{B}*}$, 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 $\mathcal{B}$: $C(\mathrm{P}^{U^*}) \le C^{\mathcal{B}*}$. Via $C^* \le C(\mathrm{P}^{U^*}) \le C^{\mathcal{B}*}$, 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 $\mathcal{B}$. Indeed observe that

$$I^{\mathcal{B}}_{t+\tau} = I^{\mathcal{B}}_t + q[t+1, t+\tau] - D[t, t+\tau-1] = S - D[t, t+\tau-1],$$

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

We will first show that $\mathbb{E}[L_t] \le \mathbb{E}[B_t]$ for any period $t \ge \tau$ when $\mathcal{L}$ and $\mathcal{B}$ operate under the same PIL-policy with level $U \ge 0$. 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.3.1.

LEMMA 4. *Let X and Y be random variables with joint distribution function $F(x,y)$, $-\infty < x, y < \infty$. Then $\mathbb{E}[(X+Y)^+] = \int_{-\infty}^{\infty} \mathbb{P}(X \ge z, Y \ge -z) dz$.*

Define the random variables

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

for $S \ge \tau\mu$. 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 $U = S - \tau\mu \ge 0$. Our aim will be to prove that $\mathbb{E}[L_\tau] \le \mathbb{E}[B_\tau]$. To this end, we first prove that X and Y are *associated* random variables (cf. Esary et al. (1967)).

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

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

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

Lemma 5. *X and Y are associated random variables.*

Proof. The random variables $\mathbf{D} = (D_0, \ldots, D_\tau)$ 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).

Observe that $J_t = (I_t - D_t)^+$ and $L_t = (D_t - I_t)^+$ so that $L_t - J_t = D_t - I_t = D_t - J_{t-1} + q_t$. This can be written as

$$J_t = J_{t-1} + q_t - D_t + L_t. \tag{17}$$

Iterating expression (17) and using $I_0 = J_{-1} + q_0$ yields

$$J_t = I_0 + q[1,t] - D[0,t] + L[0,t]. \tag{18}$$

Now assume that $L_t > 0$ (so that $L_t = D_t - I_t$) and use (18) to find

$$L_t = D_t - J_{t-1} - q_t = D_t - I_0 - q[1, t-1] + D[0, t-1] - L[0,t]. \tag{19}$$

Rearranging (19), we obtain (still under the assumption $L_t > 0$)

$$L[0,t] = D[0,t] - I_0 - q[1, t-1]. \tag{20}$$

The assumption that $L_t > 0$ can be removed by observing that

$$L[0,t] = \max_{k \in \{0,\ldots,t\}} (D[0,k] - I_0 - q[1, k-1])^+ . \tag{21}$$

Next for $Y = f(\mathbf{D})$ we have

$$Y = L[0, \tau-1] - \mathbb{E}[L[0, \tau-1]] = \max_{k \in \{0,\ldots,\tau-1\}} (D[0,k] - I_0 - q[1, k-1])^+ - \mathbb{E}[L[0, \tau-1]], \tag{22}$$

which is clearly non-decreasing in each D_i, $i \in \{1, \ldots, \tau\}$. Finally observe that

$$X = g(\mathbf{D}) = D[0,\tau] - S - \max_{k \in \{0,\ldots,\tau-1\}} (D[0,k] - I_0 - q[1,k-1])^+ + \mathbb{E}[L[0,\tau-1]]. \tag{23}$$

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

□

Lemma 6. *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 that $\mathbb{E}[B_\tau] \geq \mathbb{E}[L_\tau]$ using only that q_τ 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

$$\begin{aligned}
\mathbb{E}[B_\tau] &= \mathbb{E}\left[(X+Y)^+\right] \\
&= \int_{z=-\infty}^{\infty} \mathbb{P}(X \geq z, Y \geq -z) dz \qquad (24)\\
&\geq \int_{z=-\infty}^{\infty} \mathbb{P}(X \geq z)\mathbb{P}(Y \geq -z) dz \qquad (25)\\
&= \int_{z=-\infty}^{\infty} \mathbb{P}(\tilde{X} \geq z)\mathbb{P}(\tilde{Y} \geq -z) dz = \mathbb{E}\left[(\tilde{X}+\tilde{Y})^+\right],
\end{aligned}$$

where (24) follows from Lemma 4 and (25) follows from Theorem 5.1 of Esary et al. (1967). Now continuing and using that $\tilde{X}$ and $\tilde{Y}$ are independent, we find

$$\begin{aligned}
\mathbb{E}[B_\tau] &\geq \mathbb{E}\left[(\tilde{X}+\tilde{Y})^+\right] \\
&= \int_{x\in\mathbb{R}} \mathbb{E}_{\tilde{Y}}\left[(x+\tilde{Y})^+\right] dF_{\tilde{X}}(x) \\
&\geq \int_{x\in\mathbb{R}} x^+ dF_{\tilde{X}}(x) \qquad (26)\\
&= \mathbb{E}[\tilde{X}^+] = \mathbb{E}[X^+] = \mathbb{E}[L_\tau].
\end{aligned}$$

Inequality (26) holds because $\mathbb{E}_{\tilde{Y}}\left[(x+\tilde{Y})^+\right] \geq x + \mathbb{E}[\tilde{Y}] = x + \mathbb{E}[Y] = x$ and $\mathbb{E}_{\tilde{Y}}\left[(x+\tilde{Y})^+\right] \geq 0$ which together imply $\mathbb{E}_{\tilde{Y}}\left[(x+\tilde{Y})^+\right] \geq x^+$. □

LEMMA 7. *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 that $\mathbb{E}[c_\tau] \leq \mathbb{E}[c_\tau^{\mathcal{B}}]$ using only that q_τ 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_τ is given by:

$$I_\tau = I_0 + q[1,\tau] - D[0,\tau-1] + L[0,\tau-1]. \tag{27}$$

Under a PIL-policy with projected inventory $S - \tau\mathbb{E}[D]$, system $\mathcal{L}$ will choose q_τ such that $\mathbb{E}[I_\tau] = S - \tau\mathbb{E}[D]$. Using (27) and solving for q_τ yields that

$$q_\tau = S - I_0 - q[1,\tau-1] - \mathbb{E}[L[0,\tau-1]]. \tag{28}$$

Substituting (28) back into (27) yields

$$I_\tau = S - D[0,\tau-1] + L[0,\tau-1] - \mathbb{E}[L[0,\tau-1]]. \tag{29}$$

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

$$\begin{aligned}
\mathbb{E}[c_\tau] &= h\mathbb{E}[(I_\tau - D_\tau)^+] + p\mathbb{E}[(D_\tau - I_\tau)^+] \\
&= h\mathbb{E}[I_\tau - D_\tau] + h\mathbb{E}[(D_\tau - I_\tau)^+] + p\mathbb{E}[(D_\tau - I_\tau)^+] \\
&= h\mathbb{E}\left[S - D[0,\tau] + L[0,\tau-1] - \mathbb{E}(L[0,\tau-1])\right] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\
&= h\mathbb{E}\left[S - D[0,\tau]\right] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\
&= h\mathbb{E}\left[(S - D[0,\tau])^+\right] - h\mathbb{E}\left[(D[0,\tau] - S)^+\right] + h\mathbb{E}[L_\tau] + p\mathbb{E}[L_\tau] \\
&\leq h\mathbb{E}\left[(S - D[0,\tau])^+\right] + p\mathbb{E}\left[(D[0,\tau] - S)^+\right] = \mathbb{E}[c_\tau^{\mathcal{B}}] = C^{\mathcal{B}*}.
\end{aligned} \tag{30}$$

The second and fourth equality follow from using the identity $x = x^+ + (-x)^+$, and the inequality follows from applying Lemma 6 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 and $\mathcal{L}$ 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*

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

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: $S^*\ge\tau\mu$ and $S^*<\tau\mu$.

1. Case $S^*\ge\tau\mu$: Without loss of generality let period 0 be the first period in which it is possible to place an order to attain $U=S^*-\tau\mu\ge 0$. By Lemma 7 it holds that $\mathbb{E}[c_t]\le\mathbb{E}[c_t^{\mathcal{B}}]=C^{\mathcal{B}*}$ for all $t\ge\tau$ which implies that $C(\mathrm{P}^{S^*-\tau\mu})\le C^{\mathcal{B}^*}$.
2. Case $S^*<\tau\mu$: Observe first that $C(\mathrm{P}^0)=p\mu$ becasuse under $U=0$ there is no inventory and all demand is lost. We now have $C^{\mathcal{B}*}=C^{\mathcal{B}}(\mathrm{P}^{S^*-\tau\mu})=p\mathbb{E}[(D[0,\tau]-S^*)^+]+h\mathbb{E}[(S^*-D[0,\tau])^+]\ge p\mathbb{E}[D[0,\tau]-S^*]>p\mu=C(\mathrm{P}^0)$, where the strict inequality holds because $0\le S^*<\tau\mu$.

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

5.2. Asymptotic optimality as $p\to\infty$

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* $\lim_{x\to\infty}\mathbb{E}[D[0,\tau]-x \mid D[0,\tau]>x]/x=0$.

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 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:*

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

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

6. Numerical Results

The PIL policy is asymptotically optimal for large p and for large τ if D has an exponential distribution. Numerically, it appears that the PIL policy has good performance for large τ also for other demand distributions as Figure 1 illustrates. We benchmark the performance of the PIL policy against other policies including the optimal policy for the standard test-bed of Zipkin (2008a). We also benchmark the PIL policy against the best base-stock policy and best constant order policy on another test-bed which features higher average demands and longer leadtimes. We describe and give results for these test-beds in Sections 6.1 and 6.2.

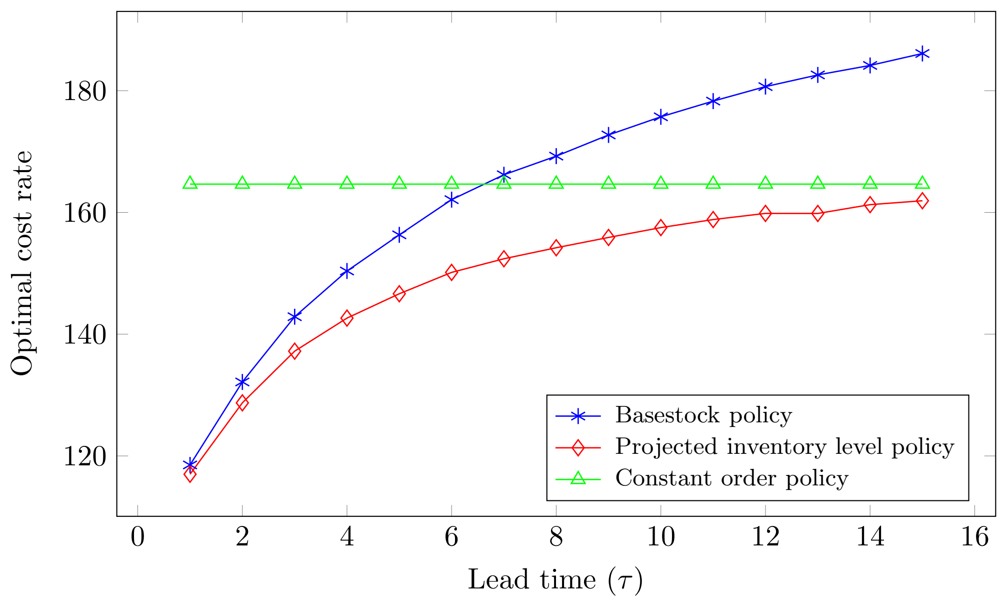

Figure 1 Cost rate as a function of lead time when $h = 1$, $p = 9$ and demand is negative binomial with 5 required successes and success probability 1/21

6.1. 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 (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:

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

The capped base-stock policy is formally defined as

$$\mathrm{R}^{S,r}(\mathbf{x}) := \min\{\mathrm{B}^S(\mathbf{x}), \mathrm{C}^r(\mathbf{x})\}.$$

Table 1 reports the performance of these policies. 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.

6.2. 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 practice. In all these instances, demand has a negative binomial distribution and the mean demand is 100 and $h = 1$. We then 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 $\sqrt{\mathbf{Var}[D]/(\mathbb{E}[D])^2} \in$

Table 1 Comparison of policies on Zipkin's test-bed

		Poisson demand				Geometric deman			
		Lead-time τ				Lead-time τ			
Penalty per lost sale	Policy	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			

$\{0.15, 0.25, 0.5, 1, 1.5, 2.0\}$ for a total of 216 instances. For each of these instances we determine the best PIL, base-stock and COP. The PIL policy dominates both the base-stock and COP policy for all instances in the test-bed. Therefore we report gap of the COP and 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 (3.68% and 41.35%). The maximum percentage gaps are strikingly large at 19.81% and 251.32% for this test-bed. This is further evidence that the PIL policy is a superior candidate for application in practical settings.

7. 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. 2020a), dual sourcing systems (e.g. Xin and Goldberg 2018), and system with independent (overtaking) lead-times (e.g. Stolyar and Wang 2019).

Acknowledgments

The authors thank Ton de Kok and Ivo Adan for stimulating discussions. The second author thanks the Netherlands Foundation for Scientific Research for funding this research.

References

Agrawal, Shipra, Randy Jia. 2019. Learning in structured mdps with convex cost functions: Improved regret bounds for inventory management. *arXiv preprint arXiv:1905.04337* .

Bai, Xingyu, Xin Chen, Menglong Li, Alexander Stolyar. 2020. Asymptotic optimality of semi-open-loop policies in markov decision processes with large lead times. *Available at SSRN* .

Table 2 Comparison of policy performance on large test-bed

	Average %-gap with best PIL		Maximum %-gap with best PIL	
Coefficient of variation of demand	BS	COP	BS	COP
0.15	5.09	61.04	19.81	251.32
0.25	4.95	54.86	18.51	230.43
0.5	4.60	46.16	15.83	181.27
1	3.52	35.83	10.51	141.33
1.5	2.36	27.72	7.07	113.89
2	1.55	22.51	4.82	93.01
Lead time (τ)				
1	1.23	65.52	5.76	251.32
2	2.39	49.95	10.18	192.43
3	3.37	40.90	13.34	155.29
4	4.26	34.77	15.92	137.58
5	5.04	30.11	17.75	116.69
6	5.76	26.88	19.81	113.63
Penalty cost (p)				
1	8.21	0.84	19.81	6.42
4	6.53	6.65	13.79	26.72
9	4.09	17.14	8.24	53.97
19	2.13	34.74	4.30	93.46
49	0.77	74.06	1.54	171.70
99	0.33	114.70	0.67	251.32
Total	3.68	41.35	19.81	251.32

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.

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

Bijvank, M., I.F.A. Vis. 2011. Lost-sales inventory theory: A review. *European Journal of Operational Research* **215** 1–13.

Bu, Jinzhi, Xiting Gong, Xiuli Chao. 2020a. Asymptotic optimality of base-stock policies for perishable inventory systems. *Available at SSRN* .

Bu, Jinzhi, Xiting Gong, Dacheng Yao. 2020b. 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, David A, Martin I Reiman, Qiong Wang. 2019. A survey of recent progress in the asymptotic analysis of inventory systems. `https://people.orie.cornell.edu/dag369/ftp/Asymptotic_Inventory_Survey_Goldberg_Reiman_Wang.pdf`. Accessed: 2020-12-16.

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** 7–11.

Song, Jing-Sheng. 1998. On the order fill rate in a multi-item, base-stock inventory system. *Operations Research* **46**(6) 831–845.

Stolyar, Alexander, Qiong Wang. 2019. Exploiting random lead times for significant inventory cost savings. *arXiv preprint arXiv:1801.02646* .

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 Fourse 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 Jaarsveld, Willem. 2020. Model-based controlled learning of mdp policies with an application to lost-sales inventory control. *arXiv preprint arXiv:2011.15122* .

Xin, Linwei. 2019. Understanding the performance of capped base-stock policies in lost-sales inventory models. *Available at SSRN 3357241* .

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 $\mathbb{E}[J_{\tau-1}] \leq U$ so that it is possible to place a non-negative order in period 0 to attain $\mathbb{E}[I_\tau] = U$. Since $\mathbb{E}[J_\tau \mid D_0 = 0]$ denotes the projected inventory level as seen in period 1 before an order is placed and when no demand occurred in period 0, it suffices to show that $\mathbb{E}[J_\tau \mid D_0 = 0] \leq \mathbb{E}[I_\tau] = U$.

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

$$(G(I_0) - D)^+ \leq G((I_0 - D)^+). \tag{EC.1}$$

It follows from (EC.1) that $\mathbb{E}[J_\tau \mid D_0 = 0] = \mathbb{E}[(G(I_0) - D)^+] \leq \mathbb{E}[G((I_0 - D)^+)] = \mathbb{E}[I_\tau]$.

EC.2. Proof of proposition 1

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 $\lim_{t\to\infty} \mathbb{E}[J_t]$ 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

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

We can now directly verify that inserting (7) and (EC.2) into the right hand side of (6) again equals (7) by using integration by parts and tedious but otherwise straightforward algebra:

$$\begin{aligned}
&\mathbb{E}_D\left[h(x-D)^+ + p(D-x)^+ + \mathcal{H}^r((x-D)^+ + r)\right] - g^r \\
&= p\mathbb{E}\left[(D-x)^+\right] + h\mathbb{E}\left[(X-D)^+\right] + \int_0^x \frac{\mathcal{H}^r(x-y+r)}{\mu} e^{-y/\mu}\,\mathrm{d}y + \mathcal{H}^r(r)e^{-x/\mu} - g^r \\
&= p\mu e^{-x/\mu} + h(x - \mu + \mu e^{-x/\mu}) + e^{-x\mu}x\left(\frac{h(3r^2+3rx+x^2)}{6\mu(\mu-r)} + \frac{pr(2r+x)}{2\mu(\mu-r)} - \frac{p/(2r+x)}{2(\mu-r)}\right) \quad \text{(EC.3)}\\
&\quad + e^{-x/\mu}\left(\frac{hr^2}{2(\mu-r)} - pr^2\right) - p(\mu-r) - h\frac{r^2}{2(\mu-r)} \\
&= \frac{h}{2(\mu-r)}x^2 - px = \mathcal{H}^r(x). \quad \text{(EC.4)}
\end{aligned}$$

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 3

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

$$\mathcal{H}^r(I_{t_1})=\mathbb{E}_{D_{t_1},\dots,D_{t_2}}\left[c[t_1,t_2](\mathrm{C}^r)+\mathcal{H}^r(I_{t_2+1})|I_{t_1}\right]-(t_2+1-t_1)g^r. \tag{EC.5}$$

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 $t_2\ge t_1$; we will show it holds also for t_2+1. Use (6) to conclude that

$$\begin{aligned}\mathbb{E}_{D_{t_1},\dots,D_{t_2}}\left[\mathcal{H}^r(I_{t_2+1})\mid I_{t_1}\right]&=\mathbb{E}_{D_{t_1},\dots,D_{t_2}}\left[\mathbb{E}_{D_{t_2+1}}[c_{t_2+1}+\mathcal{H}^r((I_{t_2+1}-D_{t_2+1})^++r)-g^r|I_{t_2+1}]\Big|I_{t_1}\right]\\&=\mathbb{E}_{D_{t_1},\dots,D_{t_2}}\left[\mathbb{E}_{D_{t_2+1}}\left[c_{t_2+1}+\mathcal{H}^r(I_{t_2+2})\mid I_{t_2+1}\right]\Big|I_{t_1}\right]-g^r\\&=\mathbb{E}_{D_{t_1},\dots,D_{t_2},D_{t_2+1}}\left[c_{t_2+1}+\mathcal{H}^r(I_{t_2+2})\mid I_{t_1}\right]-g^r.\end{aligned} \tag{EC.6}$$

Now substitution of (EC.6) back into the induction hypothesis (EC.5) yields:

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

By induction, this shows that (EC.5) holds for all $t_2\ge t_1$. $\square$

EC.3.1. Proof of Lemma 4

Proof. We have

$$\mathbb{E}[(X+Y)^+]=\int_{x+y\ge 0}(x+y)dF(x,y)=\int_{x=-\infty}^{\infty}\int_{y=-x}^{\infty}(x+y)dF(x,y). \tag{EC.8}$$

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

$$x+y=\int_{z=-y}^{x}dz. \tag{EC.9}$$

Substitution of (EC.9) into (EC.8) yields

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