OPERATIONS RESEARCH
DOI 10.1287/opre.1080.0620ec
pp. ec1–ec10

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"A Single-Unit Decomposition Approach to Multi-Echelon Inventory Systems" by Alp Muharremoglu and John N. Tsitsiklis, *Operations Research*, DOI 10.1287/opre.1080.0620

Appendix for "A Single-Unit Decomposition Approach to Multi-Echelon Inventory Systems"

by Alp Muharremoglu [1] and John N. Tsitsiklis[2]

September 2007

This Appendix is intended to supplement the paper and contains references to parts of it. References that start with "A" (e.g. Proposition A3 or Section A2) are in this Appendix, references that do not start with "A" (e.g. Proposition 3 or Section 2) are in the paper. This Appendix contains the proofs of results of Sections 5 and 6 along with supporting results.

A1 Proofs — Infinite horizon results common to both discounted cost and average cost criteria

Proof of Lemma 5.1

For any given initial state (s, z, y) for a subproblem, the number of possible future states is finite; this is because y and z cannot increase. Therefore, general results for finite-state Markov decision problems apply. When $\alpha < 1$, the convergence of $\hat{J}^*_T(s, z, y)$ to $\hat{J}^*_\infty(s, z, y)$ is immediate. When $\alpha = 1$, we have a "stochastic shortest path problem," (Bertsekas 1995) and any "improper" policy (that is, any policy that is not guaranteed to eventually deliver the unit to the customer) incurs infinite cost, due to eternal backlogging. Under this condition, the claimed convergence is again known to hold. □

Lemma A1.1. *If $u \in \hat{U}^*_t(s, z, y)$ for infinitely many choices of t, then $u \in \hat{U}^*_\infty(s, z, y)$.*

Proof. The lemma is rather elementary and we only sketch the argument. The optimality of a particular decision for a certain time horizon t can be expressed in terms of an associated Bellman equation. By taking the limit in that Bellman equation as t goes to infinity, we recover a condition that asserts optimality of the same decision for an infinite horizon problem. □

Lemma A1.2. *For every (s, z, y), with $z \in A'$, if $\hat{U}^*_\infty(s, z, y) = \{1\}$, then $1 \in \hat{U}^*_\infty(s, z, y')$, for every $y' < y$.*

Proof. If $\hat{U}^*_\infty(s, z, y) = \{1\}$, Lemma A1.1 implies that there is a $t' > 0$ such that $\hat{U}^*_t(s, z, y) = \{1\}$ for all $t > t'$. Then, by Lemma 4.1, we have $1 \in \hat{U}^*_t(s, z, y')$ for every $t > t'$ and $y' < y$. Hence, by Lemma A1.1, $1 \in \hat{U}^*_\infty(s, z, y')$ for every $y' < y$. □

[1]Graduate School of Business, Columbia University, New York, NY, 10027; e-mail: `alp2101@columbia.edu`
[2]Laboratory for Information and Decision Systems and Operations Research Center, M.I.T, Cambridge, MA, 02139; e-mail: `jnt@mit.edu`

Proof of Proposition 5.1

The argument is identical to the one in the proof of Prop.4.4, using $\hat{U}^*_\infty(s,z,y)$ in place of $\hat{U}^*_k(s,z,y)$, and by invoking Lemma A1.2 in place of Lemma 4.1. □

Lemma A1.3. *There exists a scalar $C_{\max}$ such that*

$$\hat{J}^*_T(s,N+1,y) \le \hat{J}^{\hat{\mu}^*}_T(s,N+1,y) \le \hat{J}^*_\infty(s,N+1,y) \le C_{\max},$$

for every (T,s,y), where $\hat{\mu}^*$ is the optimal subproblem policy from Prop. 5.1.

Proof. The first two inequalities are obvious, so we concentrate on the third. Note that the cost incurred by a unit-customer pair in any given single period is bounded by $b+h_{\max}$, where $h_{\max} = \max_i h_i$. Consider a policy where a unit is kept at location $N+1$ until the customer arrives, and then the unit is pushed through the system as quickly as possible (wait-push policy). Such a policy will incur a positive cost only while the unit is in transit in the system, which is at most $N+1$ periods. Thus, the infinite horizon expected cost of the wait-push policy is bounded by $C_{\max} = (N+1)\cdot(b+h_{\max})$. □

Proof of Lemma 5.2

Consider first the infinite horizon subproblem with $\alpha = 1$. Once the unit leaves stage z, a holding cost of at least $\hat{h}_{z-1}$ has to be incurred, at least until the customer arrives. Suppose that the customer position is y. Let $\tau(s,y)$ be the number of periods until the customer's arrival, given that the state of the Markov chain is currently s, and let $e(s,y) = E[\tau(s,y)]$. If the unit is released from stage z, the expected remaining cost is at least $e(s,y)\cdot\hat{h}_{z-1}$. If on the other hand, the unit is kept at z until the customer arrives and then is pushed through the system as quickly as possible, then the expected cost is at most $e(s,y)\cdot\hat{h}_z + C_{\max}$.

Clearly, for every s, $e(s,y)$ is nondecreasing in y, and diverges as y goes to infinity. Hence, there is an integer $Y^z_{\max}$ such that $e(s,y)\cdot\hat{h}_{z-1} > e(s,y)\cdot\hat{h}_z + C_{\max}$ for every $y \ge Y^z_{\max}$ and every s. Therefore, any policy that releases the unit from stage z while the position of the customer is greater than $Y^z_{\max}$ cannot be optimal. The result follows with $Y_{\max} = \max_z Y^z_{\max}$.

Consider now the infinite horizon subproblem with $\alpha < 1$. Let

$$e'(s,y) = E\left\{\sum_{t=0}^{\tau(s,y)-1} \alpha^t\right\},$$

and $f(s,y) = E[\alpha^{\tau(s,y)}]$. If the unit is released from stage z, the expected remaining cost is at least $e'(s,y)\cdot\hat{h}_{z-1}$. If on the other hand, the unit is kept at z until the customer arrives and then is pushed through the system as quickly as possible, then the expected cost is at most $e'(s,y)\cdot\hat{h}_z + f(s,y)\cdot C_{\max}$. Clearly, for every s, $f(s,y)$ converges to zero as y goes to infinity. Hence, there is an integer $Y^z_{\max}$ such that $e'(s,y)\cdot\hat{h}_{z-1} > e'(s,y)\cdot\hat{h}_z + f(s,y)\cdot C_{\max}$ for every $y \ge Y^z_{\max}$ and every s. Therefore, any policy that releases the unit from stage z while the position of the customer is greater than $Y^z_{\max}$ cannot be optimal. The result follows with $Y_{\max} = \max_z Y^z_{\max}$.

We now consider the case of a finite horizon t. If the unit is released from location z, the expected remaining cost is at least $E[\min\{\tau(s,y),t\}] \cdot \hat{h}_{z-1}$. On the other hand, the wait-push policy has a cost that is at most $E[\min\{\tau(s,y),t\}] \cdot \hat{h}_z + C_{\max}$. The difference between these two terms is $E[\min\{\tau(s,y),t\}](\hat{h}_{z-1} - \hat{h}_z) - C_{\max}$. If for a given (s,z,y,t) this difference is positive, then any policy that releases the unit at this state cannot be optimal for the subproblem. For any $y > Y^z_{\max}$,

$$\begin{aligned} & E[\min\{\tau(s,y),t\}] \cdot (\hat{h}_{z-1} - \hat{h}_z) - C_{\max} \\ & \quad \geq \; E[\min\{\tau(s,Y^z_{\max}),t\}] \cdot (\hat{h}_{z-1} - \hat{h}_z) - C_{\max}. \end{aligned}$$

As t increases to infinity, the right-hand side of the above inequality converges to:

$$\begin{aligned} & E[\tau(s,Y^z_{\max})] \cdot (\hat{h}_{z-1} - \hat{h}_z) - C_{\max} \\ & \quad = \; e(s,Y^z_{\max}) \cdot (\hat{h}_{z-1} - \hat{h}_z) - C_{\max} \\ & \quad > \; 0. \end{aligned}$$

This implies that there exists some $t^z_{\max}$ such that if $y > Y^z_{\max}$ and $t > t^z_{\max}$, then the cost of releasing the unit is larger than the cost under the wait-push policy, and therefore larger than the optimal cost. This proves the result for $t > t^z_{\max}$.

Now, let us consider a horizon length $t \leq t^z_{\max}$. The cost of a policy that keeps the unit at z will be $\hat{h}_z \cdot t$ if the customer does not show up within the time horizon, and will be at most $(b + \hat{h}_z) \cdot t$ if the customer does show up within the time horizon. Let $p_t(y)$ be the probability that a customer at position y will show up within t time periods. The expected cost of a policy that keeps the unit at z throughout the time horizon is at most $(1 - p_t(y)) \cdot \hat{h}_z \cdot t + p_t(y) \cdot (b + \hat{h}_z) \cdot t$. Now, consider a policy that releases the unit from location z. The cost of such a policy is at least $(1 - p_t(y)) \cdot t \cdot \hat{h}_{z-1}$. Since for every t, $p_t(y)$ is non-increasing in y and goes to 0 as y goes to infinity, there exists some $Y^z(t)$ such that $(1 - p_t(y)) \cdot \hat{h}_z \cdot t + p_t(y) \cdot (b + \hat{h}_z) \cdot t < (1 - p_t(y)) \cdot \hat{h}_{z-1} \cdot t$ for every $y > Y^z(t)$. This means that for any t, if $y > Y^z(t)$, then $\hat{U}^*_t(s,z,y) = \{0\}$.

The desired result follows by setting

$$Y_{\max} = \max_{z \in A'} \max \left\{ Y^z_{\max}, \max_{t < t^z_{\max}} Y^z(t) \right\}.$$

$\square$

Proof of Proposition 5.2

By definition, μ^* is decoupled. By Prop. 5.1, $\hat{\mu}^*$ is a monotonically nonincreasing function of y. Since the index i of a unit is monotonic in the position y of the corresponding customer (i.e., $i \leq j$ if and only if $y^i_t \leq y^j_t$, for all t), policy μ^* releases a unit no later than units with larger indices. It follows that μ^* is a monotonic policy as well. The rest of the argument is identical to the proof of Prop. 3.1. □

A2 Proofs — Discounted Cost

Proof of Theorem 5.1

For all monotonic states x, the policy μ^* attains the lower bound and is therefore M-optimal. By Proposition 5.2, there exists a state dependent echelon base stock policy that agrees with the monotonic and decoupled policy μ^* at every monotonic state. Therefore, this state dependent echelon base stock policy is also M-optimal. By a similar argument as in Theorem 4.1, we establish that state dependent echelon base stock policies are not only M-optimal, but optimal. □

A3 Proofs — Average Cost

Lemma A3.1. *Consider the finite horizon subproblem. There exists an integer K such that if the customer arrives K periods or more before the end of the horizon, then under any optimal policy the unit will be given to the customer before the end of the horizon.*

Proof. First, consider the infinite horizon subproblem where the customer has already arrived but the unit is not given to the customer yet. In this case, moving the unit as quickly as possible is the unique optimal control, since waiting at a certain location for one period does nothing but add one extra period of holding and backorder cost. Hence, $\hat{U}^*_\infty(s, z, 1) = \{1\}$ for every s and $z > 0$. Now fix some s and $z > 0$. By Lemma A1.2, we cannot have $0 \in \hat{U}^*_t(s, z, 1)$ for infinitely many choices of t and therefore there exists some $t'(s, z)$ such that $\hat{U}^*_t(s, z, 1) = \{1\}$ for every $t > t'(s, z)$. Let $K = N + 2 + \max_{s,z} t'(s, z)$. Now, consider the T-horizon subproblem and suppose that a customer arrives at time k where $T - k \geq K$. In this case, the optimal decision is to move the unit through the system as quickly as possible, and give it to the customer before the end of the horizon. □

Lemma A3.2. *Fix some $\epsilon > 0$. For every s and for every y such that $y \leq \left(\bar{d} - \epsilon\right) \cdot T$, we have*

$$\left| \hat{J}^*_T\left(s, N+1, y\right) - \hat{J}^*_\infty\left(s, N+1, y\right) \right| \leq f\left(T, s\right)$$

for some $f : \mathbb{N} \times S \mapsto \mathbb{R}$ such that $\lim_{T\to\infty} f\left(T, s\right) = 0$.

Proof. Note that $\hat{J}^*_T\left(s, N+1, y\right) \leq \hat{J}^*_\infty\left(s, N+1, y\right)$, since having more time periods can only increase the costs. We will next establish an inequality in the reverse direction.

Given a time horizon, consider the following non-stationary policy for the infinite horizon subproblem with initial state $(s, N+1, y)$: in the first T periods, employ an optimal policy for the T-horizon problem, and then employ the wait-push policy. At the end of period T, the unit is either at location 0 (given to the customer), or at a location greater than 0 but less than $N+1$ (in the system), or at location $N+1$ (at the supplier). If the unit is given to the customer, there is no more cost. If the unit is in the system, then the position of the customer is at most $Y_{\max}$, by Lemma 5.2; hence, the remaining cost can be bounded by some v that does not depend on y. If the unit is at the supplier, then the cost of the wait-push policy is bounded by $C_{\max}$, by Lemma A1.3. Hence, the cost of this combined policy is bounded by

$$\hat{J}_T^*(s, N+1, y) + \text{Prob}\{\text{Unit not given to the customer by time } T\} \cdot (C_{\max} + v).$$

Moreover, this cost has to be at least as large as the optimal infinite horizon cost. So, we have:

$$\begin{aligned} &\hat{J}_\infty^*(s, N+1, y) - \hat{J}_T^*(s, N+1, y) \\ &\leq \text{Prob}\{\text{Unit not given to the customer by time } T\} \cdot (C_{\max} + v). \end{aligned}$$

Let $d^{T-K}(s)$ be a random variable denoting the sum of the demands in $T-K$ periods, starting with a period where the Markov chain is in state s. If the unit is not given to the customer by time T, then by Lemma A3.1, the customer has not arrived by time $T-K$, that is, $d^{T-K}(s) < y-1$. Hence,

$$\hat{J}_\infty^*(s, N+1, y) - \hat{J}_T^*(s, N+1, y) \leq \text{Prob}\left\{d^{T-K}(s) < (\bar{d}-\epsilon)T\right\} \cdot (C_{\max} + v).$$

For $T > K$, let $f(T, s)$ be the right hand side of the above inequality. It remains to show that $\lim_{T\to\infty} f(T, s) = 0$.

Indeed,

$$\text{Prob}\left\{d^{T-K}(s) < (\bar{d}-\epsilon)T\right\} = \text{Prob}\left\{\frac{d^{T-K}(s)}{T} < \bar{d}-\epsilon\right\}.$$

As $T \to \infty$, by the law of large numbers for Markov reward processes, $d^{T-K}(s)/T$ converges to $\bar{d}$ almost surely, and therefore, in probability. Therefore, the probability we are considering converges to zero, and so does $f(T, s)$. □

Proof of Theorem 5.2

We first show that the policy μ^* from Prop. 5.2 is M-optimal. Since μ^* is a decoupled policy, we have for every finite horizon T,

$$V_T^{\mu^*}(x) = \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, z^i, y^i\right).$$

Since μ^* is also monotonic, Proposition 4.2(b) yields

$$J_T^{\mu^*}(x) = V_T^{\mu^*}(x) = \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, z^i, y^i\right),$$

for every monotonic state x. Then, by the definition of the infinite horizon average cost,

$$\lambda^{\mu^*}(x) = \limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, z^i, y^i\right),$$

for every monotonic state x.

Let us fix a monotonic initial state $x = \left\{s, \left(z^1, y^1\right), \left(z^2, y^2\right), \ldots\right\}$. If there is an infinite number of units in locations other than $N+1$, the optimal average cost $\lambda^*(x)$ is infinite, and there is nothing to prove. We can therefore assume there there is a finite number k of units in locations $1, \ldots, N$. Let ℓ be the number of units whose corresponding customers have already arrived, so that $y^i = i - l + 1$ for $i \geq l$. We have

$$\begin{aligned}
\lambda^{\mu^*}(x) &= \limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, z^i, y^i\right) \\
&= \limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{k+\ell} \hat{J}_T^{\hat{\mu}^*}\left(s, z^i, y^i\right) + \limsup_{T\to\infty} \frac{1}{T} \sum_{i=k+\ell+1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i-\ell+1\right) \\
&= \limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right).
\end{aligned}$$

Let $\epsilon > 0$ be a constant less than $\bar{d}/2$. We decompose the above expression as follows:

$$\begin{aligned}
&\limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right) \\
&= \limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\lceil \epsilon\cdot T\rceil} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right) + \limsup_{T\to\infty} \frac{1}{T} \sum_{i=\lceil \epsilon\cdot T\rceil+1}^{\left\lceil\left(\bar{d}-\epsilon\right)\cdot T\right\rceil} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right) \\
&\quad + \limsup_{T\to\infty} \frac{1}{T} \sum_{i=\left\lceil\left(\bar{d}-\epsilon\right)\cdot T\right\rceil+1}^{\left\lceil\left(\bar{d}+\epsilon\right)\cdot T\right\rceil} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right) + \limsup_{T\to\infty} \frac{1}{T} \sum_{i=\left\lceil\left(\bar{d}+\epsilon\right)\cdot T\right\rceil+1}^{\infty} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right).
\end{aligned} \tag{A1}$$

We will show that the first and the third terms in the above sum go to zero as $\epsilon \to 0$, and that the fourth term is equal to zero.

Using Lemma A1.3, the first term satisfies

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{i=1}^{\lceil \epsilon T\rceil} \hat{J}_T^{\hat{\mu}^*}\left(s, N+1, i\right) \leq \limsup_{T\to\infty} \frac{1}{T} \epsilon T C_{\max} \underset{\epsilon\to 0}{\longrightarrow} 0.$$

Similarly, for the third term,

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil(\bar{d}-\epsilon)\cdot T+1\rceil}^{\lceil(\bar{d}+\epsilon)\cdot T\rceil}\hat{J}_T^{\hat{\mu}^*}(s,N+1,i)\leq\limsup_{T\to\infty}\frac{1}{T}2\epsilon TC_{\max}\underset{\epsilon\to0}{\longrightarrow}0.$$

To get the result for the fourth term, consider a unit i and its corresponding customer. After an interval of T periods, the position of this customer will be $\left(i-d^T(s)\right)^++1$, where $d^T(s)$ is the random variable denoting the sum of demands in T periods, starting from a period with the Markov chain in state s (assuming that the customer has not received the unit within the interval). By Lemma 5.2, if $\left(i-d^T(s)\right)^++1>Y_{\max}$, unit i will not be released from location $N+1$ and this unit-customer pair will have a cost of 0 during the T-step horizon. For any unit that is released from location $N+1$, the expected cost can be at most $C_{\max}$. Therefore,

$$\begin{aligned}&\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil(\bar{d}+\epsilon)\cdot T\rceil+1}^{\infty}\hat{J}_T^{\hat{\mu}^*}(s,N+1,i)\\ \leq\;&\limsup_{T\to\infty}\frac{1}{T}\cdot C_{\max}\cdot E\left\{\left(d^T(s)+Y_{\max}-\lceil(\bar{d}+\epsilon)\cdot T\rceil-1\right)^+\right\}=0,\end{aligned}$$

using the law of large numbers. We have therefore established that only the second term in the right hand side of Eq. (A1) remains positive as $\epsilon\downarrow 0$, and

$$\begin{aligned}\lambda^{\mu^*}(x)&=\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil\epsilon\cdot T\rceil+1}^{\lceil(\bar{d}-\epsilon)\cdot T\rceil}\hat{J}_T^{\hat{\mu}^*}(s,N+1,i)+f(\epsilon)\\&\leq\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil\epsilon\cdot T\rceil+1}^{\lceil(\bar{d}-\epsilon)\cdot T\rceil}\hat{J}_\infty^{\hat{\mu}^*}(s,N+1,i)+f(\epsilon),\end{aligned}$$

for some function f that satisfies $\lim_{\epsilon\downarrow0}f(\epsilon)=0$.

We now use Eq. (2), to obtain

$$\begin{aligned}\lambda^*(x)&\geq\limsup_{T\to\infty}\frac{1}{T}\sum_{i=1}^{\infty}\hat{J}_T^*\left(s,z^i,y^i\right)\\&\geq\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil\epsilon\cdot T\rceil+1}^{\lceil(\bar{d}-\epsilon)\cdot T\rceil}\hat{J}_T^*(s,N+1,i)\\&\geq\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil\epsilon\cdot T\rceil+1}^{\lceil(\bar{d}-\epsilon)\cdot T\rceil}\hat{J}_\infty^*(s,N+1,i)-\limsup_{T\to\infty}\frac{1}{T}\bar{d}Tf(T,s)\\&=\limsup_{T\to\infty}\frac{1}{T}\sum_{i=\lceil\epsilon\cdot T\rceil+1}^{\lceil(\bar{d}-\epsilon)\cdot T\rceil}\hat{J}_\infty^*(s,N+1,i),\end{aligned}$$

where the last inequality uses Lemma A3.2. By comparing the above two inequalities, and using the fact $\hat{J}^*_\infty = \hat{J}^{\hat{\mu}^*}_\infty$ (optimality of $\hat{\mu}^*$ for the infinite horizon subproblem), we obtain

$$\lambda^{\mu^*}(x) \leq \lambda^*(x) + f(\epsilon).$$

By taking the limit as ϵ decreases to zero, we obtain $\lambda^{\mu^*}(x) \leq \lambda^*(x)$, which establishes the M-optimality of μ^*.

By Prop. 5.2, μ^* agrees with a state dependent echelon base stock policy at monotonic states, establishing the M-optimality of state dependent echelon base stock policies. Following the argument in Theorem 4.1, state dependent echelon base stock policies are not only M-optimal, but optimal. □

Proof of Proposition 5.3

a) Let $F_s(r)$ be the probability mass function of the distribution of the demand when the Markov chain is in state s. Let $P_{i,j} = \mathrm{P}(s_{t+1} = j \mid s_t = i)$ be the transition probabilities of the Markov chain s_t. By Lemma 5.2, there exists a positive number $Y_{\max}$ such that if $y > Y_{\max}$, the optimal decision is to not release the unit, i.e., $u = 0$. Thus, if $y > Y_{\max}$, the dynamic programming equation for the subproblem yields

$$\hat{J}^*_\infty(s, N+1, y) = \hat{g}(s, N+1, y, 0) + \sum_{j \in S} \sum_{r=0}^{\infty} P_{s,j} \cdot F_s(r) \cdot \hat{J}^*_\infty\left(j, N+1, (y-1-r)^+ + 1\right).$$

This equation is in the form of a Markov renewal equation. Since the Markov chain s_t is irreducible and aperiodic, and since the demand process d_t is of the non-lattice type, Prop. 4.17 in Chapter 10 of Cinlar (1975) applies and shows that the solution of the Markov renewal equation converges as y goes to infinity, to a constant that does not depend on s.

b) The earlier outlined proof of Theorem 5.2 shows that

$$\begin{aligned}
\lambda^*(x) &= \lim_{\epsilon \downarrow 0} \left[\limsup_{T \to \infty} \frac{1}{T} \sum_{i=\lceil \epsilon \cdot T \rceil + 1}^{\lceil (\bar{d}-\epsilon) \cdot T \rceil} \hat{J}^*_\infty(s, N+1, i) + f(\epsilon) \right] \\
&= \lim_{\epsilon \downarrow 0} [(\bar{d} - 2\epsilon) C + f(\epsilon)] \\
&= \bar{d} C,
\end{aligned}$$

where the second equality follows from part (a). □

A4 Proofs — Algorithmic Issues

Proof of Proposition 6.1

a) By Lemma 5.2, if $y > Y_{\max}$, then $\hat{U}^*_t(s, z, y) = \{0\}$ for all s, $z \in A'$, and t. Therefore, all the base stock levels will be determined at this point and the algorithm will terminate.

b) Fix some s, $z \in A'$, and t. First, note that $y_t^*(s,z)$ is the largest y for which $\hat{U}_t^*(s,z,y) = \{1\}$, if such a y exists, and is equal to $-\infty$ if there is no y for which $\hat{U}_t^*(s,z,y) = \{1\}$. (Note that by Lemma 5.2, $\hat{U}_t^*(s,z,y)$ cannot be equal to $\{1\}$ for infinitely many y.) Then, by the arguments found in the proof of Prop. 4.4, a decoupled policy that uses a monotonic subproblem policy for each unit-customer pair with threshold levels $y_t^*(s,z)$ is M-optimal for the Main Model. By Prop. 3.1, this policy agrees with a state dependent echelon base stock policy with base stock levels $S_t^{v(z-1)}(s) = y_t^*(s,z) - 1$ at every monotonic state. By the argument in Theorem 4.1, this state dependent echelon base stock policy is not only M-optimal, but optimal.

c) Each calculation of the function $V_t(s,z,y,u)$ takes $O(\min\{Y_{\max}, D\} \cdot |S|)$ time, because for every state, the number of next possible states is of that order and there are only two possible controls. This is performed for $O(Y_{\max} \cdot T \cdot N \cdot |S|)$ times, via the nested *while* and *for* loops. Hence, the complexity of FHA is $O\Big(N \cdot Y_{\max} \cdot \min\{Y_{\max}, D\} \cdot |S|^2 \cdot T\Big)$. □

Definition A4.1. Consider an n-state Markov chain. Suppose that there is a cost for being at a given state. In addition, suppose that there is a controller that has an option to stop the Markov chain at any time, and that for each state there is a cost associated with stopping the Markov chain at that state. After the Markov chain is stopped, no more costs are incurred. The *optimal stopping problem* looks for a stopping policy that minimizes the total infinite horizon expected cost.

Lemma A4.1. *An optimal stopping problem with n states can be solved in $O(n^3)$ time.*

Proof. (Outline) Consider the policy iteration algorithm, starting with the policy that stops at every state. We claim that the policy iteration algorithm (with ties broken in favor of stopping) takes at most n iterations. Indeed, the cost-to-go of a state cannot increase in the course of the policy iteration algorithm. Thus, if the cost-to-go of not stopping becomes smaller than the cost of stopping at a given state, it remains smaller in subsequent iterations. Thus, with each policy iteration, the policy is either the same (in which case, we have an optimal policy), or the number of states at which the policy does not stop increases.

Let k_i be the number of states at which the policy changes at the i^{th} iteration. At each policy iteration, there is a policy evaluation step in which we need to solve a new system of equations with n unknowns. But the transition matrix differs from the previous one in only k_i rows (the ones where the policy changed). This is a rank k_i modification. We can use the following fact from numerical linear algebra. If A is $n \times n$ and A^{-1} is available, and if $B - A$ has rank k, then we can compute B^{-1} in $O(n^2 \cdot k)$ time. (This is possible because of the Sherman-Morrison-Woodbury formula in Golub & van Loan (1983)). Therefore, the total complexity is $O(\sum_i k_i \cdot n^2)$ and since $\sum_i k_i \cdot n^2 \leq n^3$, the result follows. □

Lemma A4.2. *Fix a pair (z,y) consisting of a unit location z and a customer position $y \leq Y_{\max}$. Suppose that $\hat{J}_\infty^*(s',z',y')$ is available for every (s',z',y') such that $z' \leq z$, $y' \leq y$, and $(z',y') \neq (z,y)$. Then, the values of $\hat{J}_\infty^*(s,z,y)$, for all $s \in S$, can be found in time $O\Big(\min\{Y_{\max}, D\} \cdot |S|^2\Big)$*

if z is an artificial stage ($z \notin A$), and in time $O\left(\min\{Y_{\max}, D\} \cdot |S|^2 + |S|^3\right)$ if z is an actual stage ($z \in A$).

Proof. Consider first the case where z is an artificial stage. Then, the location z' at the next time is guaranteed to satisfy $z' < z$. The Bellman equation for $\hat{J}^*_\infty(s, z, y)$ involves the known values of $\hat{J}^*_\infty(s', z', y')$ for the various possible next states (s', z', y'). For each $s \in S$, there are at most $2 \cdot \min\{Y_{\max}, D\} \cdot |S|$ possible next states, and the complexity estimate $O\left(\min\{Y_{\max}, D\} \cdot |S|^2\right)$ follows.

Suppose now that z corresponds to an actual stage ($z \in A$), but $z > 1$. Given a current state (s, z, y), the successor state is of the form (s', z, y) as long as the demand is zero and the decision is to not release the unit. We view a release decision as a stopping decision and a nonzero demand as a forced stopping. When we write down the Bellman equation for the various states of the form (s, z, y), for a fixed pair (z, y), it takes the form of the Bellman equation for an optimal stopping problem for a Markov chain with $|S|$ states. The transition probabilities, stopping, and continuation costs for this optimal stopping problem can be computed in time $O\left(\min\{Y_{\max}, D\} \cdot |S|^2\right)$. (This is because we have $|S|$ states of the form (s, z, y) and for each such state at most $2 \cdot \min\{Y_{\max}, D\} \cdot |S|$ possible next states.) By Lemma A4.1, the corresponding optimal stopping problem can be solved in $O\left(|S|^3\right)$ time. For the case when $z = 1$, there is no decision to release a unit, i.e., no stopping decision, and stopping occurs only via the forced stopping of a nonzero demand. Nevertheless, the same complexity of $O\left(|S|^3\right)$ applies. Hence, the claimed complexity estimate follows.

□

Proof of Proposition 6.2

(a-b) The proofs are similar to the proofs of the finite horizon versions (Proposition 6.1 parts (a) and (b)).

c) For a given (z, y) pair, the complexity of computing $\hat{J}^*_\infty(s, z, y)$ for all s is given by Lemma A4.2. This is done every time the recursion gets to (*). There are $O(N \cdot Y_{\max})$ pairs (z, y) to be considered, and $O(M \cdot Y_{\max})$ pairs for which z corresponds to an actual stage. The result follows. □

References

Bertsekas, D. (1995), *Dynamic programming and optimal control*, Athena Scientific, Belmont, MA.

Cinlar, E. (1975), *Introduction to Stochastic Processes*, Prentice-Hall Inc., Englewood Cliffs, NJ.

Golub, G. H. & Van Loan, C. F. (1983), *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD.