A Sample-Path Approach to the Optimality of Echelon Order-Up-To Policies in Serial Inventory Systems

Woonghee Tim Huh, Department of IEOR, Columbia University
Ganesh Janakiraman, Stern School of Business, New York University

May 19, 2008

Abstract

We present a new proof of the optimality of echelon order-up-to policies in serial inventory systems, first proved by Clark and Scarf. Our proof is based on a sample-path analysis as opposed to the original proof based on dynamic programming induction.

Keywords: Inventory, Multi-echelon, Sample-path analysis

*Corresponding author. Industrial Engineering and Operations Research Department, Columbia University, 500 West 120th Street, New York, NY 10027, USA. Email: huh@ieor.columbia.edu.
1 Introduction

In this paper, we provide an alternate proof of the optimality of echelon order-up-to policies in serial inventory systems, first proved in the seminal paper by Clark and Scarf [5]. Our proof is based on a sample-path analysis as opposed to the original proof based on dynamic programming induction. The crux of our proof is the following claim: for every policy that is not of the echelon order-up-to type, we can construct a policy that dominates the original policy strictly.

We now highlight the weaknesses and strengths of our approach relative to the original. In our analysis, we assume just as Clark and Scarf do that the single period cost is a sum of $N$ functions (where $N$ is the number of echelons) of the individual echelon inventory positions. While Clark and Scarf require these functions to be weakly convex, we require strict convexity. (Although this is a stricter assumption than weak convexity, note that any convex function can be approximated arbitrarily closely by a strictly convex function.) Furthermore, their proof also provides an algorithm for computing the optimal order-up-to levels whereas our analysis does not. In these aspects, our analysis is weaker than the original. On the other hand, we believe that our proof is simpler, especially for those who are not familiar with the technique of dynamic programming induction. Given that the optimality of echelon order-up-to policies is one of the most important results in inventory theory, we think that there is merit in developing new proofs that are fundamentally different from the original.

The only other alternate proofs of this result, that we are aware of, are due to [3], [4] and [6]. While elegant and important in their own right, they require the cost function to be composed of linear holding and shortage costs. (Strictly speaking, the analysis of [6] allows cost functions that can depend on the waiting time of an individual customer and the time an inventory unit spends in an echelon. However, when we restrict attention to costs that are functions of echelon inventory positions alone, only linear holding and shortage costs are allowed.) Moreover, [4] requires demands to be independently and identically distributed while [3] requires demands to be Markov-modulated; in contrast, we do not make any assumptions on the evolution of demands except that demands are unaffected by the ordering decisions. Additionally, these two papers consider the long run average cost criterion whereas our analysis holds for the finite horizon or the infinite horizon discounted cost criterion.
Recently, there have been some notable generalizations of these optimality results: for example, [2] and [7] study serial inventory systems in which each echelon has a specified replenishment frequency.

2 Problem Definition

We consider an uncapacitated two-echelon serial inventory system under periodic review. Demand occurs at the lower stage (stage 1), which orders from the upper stage (stage 2); stage 2 orders from an outside source with ample supply. The replenishment lead times in the system are deterministic, and we assume that both lead times are one period in length. We assume that demands in any two distinct periods are independent of each other. (These assumptions – that there are exactly two stages, that the lead times are one period in length, and that demands are independently distributed – are made only for the sake of exposition. Our proof extends readily to an arbitrary number of stages, arbitrary lead time lengths and any demand process whose evolution is unaffected by the policy.) We consider a finite planning horizon $T$, and periods are indexed forward by $t$. Let $(D_t | t = 1, 2, \ldots, T)$ denote the sequence of nonnegative random variables representing demand.

The sequence of events in each period $t$ is given as follows. (i) Each echelon $j \in \{1, 2\}$ receives the delivery due to arrive in period $t$, which is the same as the quantity ordered in the previous period, $q_{j-1}^j$. Any backlogged demand at stage 1 is immediately satisfied to the extent possible. (ii) The manager observes the amount of inventory in each echelon $j$. Let $x_j^t$ denote the net echelon inventory level for echelon $j$, which is the total amount of physical inventory at stages $j, \ldots, 1$ minus the amount of backordered demand at stage 1. (iii) The manager makes the ordering decision $q_j^t \geq 0$ for each stage $j = 1, \ldots, N$. We require that $q_j^t$ cannot exceed the amount of inventory available in the immediately upstream stage, which is $x_{j+1}^t - x_j^t$. (For simplicity, let $x_3^t = \infty$ for all $t$.) Let $y_j^t$ denote the after-ordering echelon inventory position for echelon $j$, i.e., $y_j^t = x_j^t + q_j^t$, which satisfies

$$y_j^t \in [x_j^t, x_j^{t+1}] .$$

(iv) Demand $D_t$ is realized and is satisfied to the extent possible. Any excess demand is backordered to the next period. The dynamics of the system can be represented as follows: $x_{t+1}^j = y_j^t - D_t$.
The expected single-period cost for period $t$, $\phi_t$, is given by a separable convex function of $(y_1^t, y_2^t)$, i.e.,

$$\phi_t(y_1^t, y_2^t) = L_1^t(y_1^t) + L_2^t(y_2^t),$$

where both $L_1^t$ and $L_2^t$ are strictly convex. We assume that each $L_j^t(y_j^t) \to \infty$ as $y_j^t \to \infty$ or $-\infty$.

The objective is to minimize the $T$-horizon expected cost $E[\sum_{t=1}^{T} \phi_t(y_1^t, y_2^t)]$. The above problem can be formulated as a dynamic program where the state and action pair in each period is given by $x_t = (x_1^t, x_2^t)$ and $y_t = (y_1^t, y_2^t)$.

An echelon order-up-to policy is an ordering policy characterized by a vector of order-up-to levels $S_t = (S_1^t, \ldots, S_N^t)$, which is defined for each period $t$. The ordering quantity at each stage is chosen to bring the echelon inventory level as close to the order-up-to level as possible. We also call this the echelon order-up-to $S$ policy. Thus, for $j \in \{1, 2\}$,

$$q_j^t = \left[ \min(S_j^t, x_t^{j+1}) - x_j^t \right]^+,$$

or equivalently $y_j^t = \max(\min(S_j^t, x_t^{j+1}), x_j^t)$.

### 3 Analysis

The following proposition provides a useful characterization of echelon order-up-to policies.

**Proposition 1.** For any $t \in \{1, \ldots, T\}$, the following statements are equivalent for the serial inventory system of Section 2:

(i) There exists an order-up-to vector $S_t = (S_1^t, S_2^t)$ such that the system follows an echelon order-up-to $S_t$ policy in period $t$.

(ii) For every pair of states $x_t^A = (x_t^{A1}, x_t^{A2})$ and $x_t^B = (x_t^{B1}, x_t^{B2})$, and the corresponding pair of actions $y_t^A = (y_t^{A1}, y_t^{A2})$ and $y_t^B = (y_t^{B1}, y_t^{B2})$, exactly one of the following three statements holds for each $j \in \{1, 2\}$:

- $y_t^{Aj} = y_t^{Bj}$;
- $y_t^{Aj} < y_t^{Bj}$ and either $y_t^{Bj} = x_t^{Bj}$ or $y_t^{Aj} = x_t^{A,j+1}$;
- $y_t^{Aj} > y_t^{Bj}$ and either $y_t^{Aj} = x_t^{A,j} = x_t^{Aj}$ or $y_t^{Bj} = x_t^{B,j+1}$.
Proof. The claim that (i) implies (ii) follows directly from the definition of an echelon order-up-to policy. We will now prove that (ii) implies (i). Let \( y_t(x_t) = (y^C_t(x_t), y^D_t(x_t)) \) denote the action (inventory position after ordering) for the given state (inventory level before ordering) \( x_t = (x^C_t, x^D_t) \). Fix \( j \in \{1, 2\} \). Recall \( y^j_t \in [x^j_t, x^{j+1}_t] \) from (1). We consider the following cases separately.

Case I. If \( y^j_t(x_t) = x^j_t \) for all \( x_t \), or \( y^j_t(x_t) = x^{j+1}_t \) for all \( x_t \), then echelon \( j \) follows the order-up-to \( S^j_t \) policy in period \( t \), where \( S^j_t = -\infty \) or \( S^j_t = \infty \), respectively.

Case II. Suppose \( y^j_t(x_t) \in \{x^j_t, x^{j+1}_t\} \) for all \( x_t \). Then, define

\[
L_j = \{x_t : x^j_t < y^j_t(x_t) = x^{j+1}_t\}, \quad R_j = \{x_t : x^j_t = y^j_t(x_t) < x^{j+1}_t\} \quad \text{and} \quad M_j = \{x_t : x^j_t = x^{j+1}_t\}.
\]

Note that for every \( j \in \{1, 2\} \), any \( x_t \) belongs to exactly one of the sets \( L_j, R_j \) and \( M_j \). Then, we have \( \hat{x}^j_t \leq \bar{x}^j_t \) for any \( \hat{x}_t \in L_j \) and \( \bar{x}_t \in R_j \). (Otherwise, we must have \( y^j_t(\hat{x}_t) > \hat{x}^j_t > \bar{x}^j_t = y^j_t(\bar{x}_t) \), which imply by (ii) that \( y^j_t(\hat{x}_t) > \hat{x}^j_t \) or \( y^j_t(\bar{x}_t) = \bar{x}^j_t \). In the first case, \( \hat{x}_t \in L_j \) implies \( \hat{x}^j_t = y^j_t(\hat{x}_t) = \bar{x}^j_t \), contradicting the definition of \( L_j \). Similarly, in the other case, \( \hat{x}_t \in R_j \) implies \( \hat{x}^j_t = y^j_t(\hat{x}_t) = \bar{x}^j_t \), contradicting the definition of \( R_j \).) Thus, there exists \( S^j_t \) such that \( \hat{x}^j_t \leq S^j_t \leq \bar{x}^j_t \) for any \( \hat{x}_t \in L_j \) and \( \bar{x}_t \in R_j \). It is now easy to verify that for every \( x_t \) in \( L_j, R_j \) and \( M_j \), the policy followed by echelon-\( j \) satisfies the definition of the order-up-to \( S^j_t \) policy.

Case III. We proceed by assuming otherwise, that is, there exists \( \hat{x}_t \) such that \( y^j_t(\hat{x}_t) \in (\hat{x}^j_t, \bar{x}^{j+1}_t) \).

Define \( S^j_t = y^j_t(\hat{x}_t) \). For any state vector \( \hat{x}_t \), the condition of this proposition implies either (a) \( y^j_t(\hat{x}_t) = \hat{y}^j_t(\hat{x}_t) \), (b) \( y^j_t(\hat{x}_t) < \hat{y}^j_t(\hat{x}_t) = \hat{x}^j_t \), or (c) \( y^j_t(\hat{x}_t) > \hat{y}^j_t(\hat{x}_t) = \hat{x}^{j+1}_t \). Thus, echelon \( j \) follows the order-up-to \( S^j_t \) policy in period \( t \).

The key idea for our sample-path approach is based on the following definition. Let \( A \) and \( B \) be any pair of systems starting from initial states \( x^A_1 = (x^A_1, x^A_2) \) and \( x^B_1 = (x^B_1, x^B_2) \), respectively. We say that a pair of systems \( (C, D) \) is the tightened pair of \( (A, B) \) if both of the following conditions are satisfied:

(i) The initial states for \( C \) and \( D \) are \( x^A_1 \) and \( x^B_1 \), respectively.

(ii) In each period \( t \), \( y^C_t = (y^C_t, y^C_t) \) and \( y^D_t = (y^D_t, y^D_t) \) are feasible actions, which solve

\[
\min_{y^C_t, y^D_t} \left\{ y^C_t - y^D_t : y^C_t + y^D_t = y^A_t + y^B_t, x^j_t \leq y^C_t \leq x^{j+1}_t, x^j_t \leq y^D_t \leq x^{j+1}_t \right\}
\]

for each \( j \in \{1, 2\} \).
The existence of the tightened pair $x^A_t$ and $x^B_t$, respectively. The second condition states that systems $C$ and $D$ have the same initial states as $A$ and $B$, respectively.

Proposition 2. Let $A$ and $B$ be a pair of systems starting from initial states $x^A_1$ and $x^B_1$, respectively, and let $(C, D)$ be the tightened pair of $(A, B)$. Then, for any $t \in \{1, \ldots, T\}$ and $j \in \{1, 2\}$,

$$\min(y^A_{Cj}, y^B_{Dj}) \leq \min(y^C_t, y^D_t) \leq \max(y^C_t, y^D_t) \leq \max(y^A_{Cj}, y^B_{Dj}).$$

(4)

Proof. If $t = 1$, then $y^C_t = y^A_t$ and $y^D_t = y^B_t$ are feasible decisions in the optimization problem defined in (3). Thus, the pair of $y^C_t$ and $y^D_t$ optimizing (3) satisfies $|y^C_t - y^D_t| \leq |y^A_t - y^B_t|$ for each $j$. From the fact $y^C_t$ and $y^D_t$ are optimal and $y^C_t + y^D_t = y^A_t + y^B_t$, it follows that (4) holds for $t = 1$. We proceed by induction with the hypothesis that feasible $(y^C_{t-1}, y^D_{t-1})$ exists and satisfies (4) for $t - 1$, where $t > 1$. Since $x^C_t = x^D_t - D_{t-1}$, it follows from (3) that, for each $j$,

$$x^A_j + x^B_j = x^C_j + x^D_j$$

(5)

and

$$\min(x^A_j, x^B_j) \leq \min(x^C_j, x^D_j) \leq \max(x^C_j, x^D_j) \leq \max(x^A_j, x^B_j).$$

(6)

Since $y^A_t$ is a feasible action in system $A$ and $y^B_t$ is a feasible action in system $B$, it follows from (1) that $y^A_t \in [x^A_j, x^A_{j+1}]$ and $y^B_t \in [x^B_j, x^B_{j+1}]$. From this observation and (6), we obtain the following inequalities:

$$\min(y^A_t, y^B_t) \leq \min(x^A_{j+1}, x^B_{j+1}) \leq \min(x^C_{j+1}, x^D_{j+1}).$$

(7)

and

$$\max(y^A_t, y^B_t) \geq \max(x^A_t, x^B_t) \geq \max(x^C_t, x^D_t).$$

(8)

Note that, in the optimization problem (3), the upper bounds on $y^C_t$ and $y^D_t$ are given by $x^C_{j+1}$ and $x^D_{j+1}$, respectively, and the lower bounds are given by $x^C_t$ and $x^D_t$, respectively. Since $(C, D)$ is the tightened pair of $(A, B)$, it can now be verified using (5), (7) and (8) that $\min(y^C_t, y^D_t) \geq \min(y^A_t, y^B_t)$ and $\max(y^C_t, y^D_t) \leq \max(y^A_t, y^B_t)$, completing the induction argument. □
If both systems $A$ and $B$ follow the echelon order-up-to policy with the same order-up-to vector, then it can easily be seen that the tightened pair of $(A, B)$ is the same as $(A, B)$ itself. Otherwise, we show in Proposition 3 that the tightened pair of these systems strictly improves the combined performance. More specifically, we show that a policy which is not of the echelon order-up-to type in period 1 can be dominated strictly by another policy from at least one starting state. Here, the choice of period 1 as the period of interest is without any loss of generality.

**Proposition 3.** Consider a policy that is not of the echelon order-up-to type in period 1. Then, there exists a pair of states $(x^A_1, x^B_1)$ such that the following statements are true, where $A$ and $B$ denote a pair of systems starting from $(x^A_1, x^B_1)$, and $(C, D)$ is the tightened pair of $(A, B)$:

(a) There exists $j \in \{1, 2\}$ such that
\[
\min(y^A_1, y^B_1) < \min(y^C_1, y^D_1) \leq \max(y^C_1, y^D_1) < \max(y^A_1, y^B_1).
\]

(b) For any $T^o \geq 1$,
\[
\sum_{t=1}^{T^o} \phi_t(y^A_1, y^A_2) + \sum_{t=1}^{T^o} \phi_t(y^B_1, y^B_2) > \sum_{t=1}^{T^o} \phi_t(y^C_1, y^C_2) + \sum_{t=1}^{T^o} \phi_t(y^D_1, y^D_2). \tag{9}
\]

**Proof.** Since systems $A$ and $B$ do not follow an echelon order-up-to policy in period 1, Proposition 1 implies the existence of two state-action pairs $(x^A_1, y^A_1)$ and $(x^B_1, y^B_1)$ that violate the conclusions of Proposition 1 (ii), i.e., there exists some $j \in \{1, 2\}$ such that
\[
y^A_1 < y^B_1, \quad y^A_1 \neq x^B_1, \quad \text{and} \quad y^A_1 \neq x^A_{j+1}, \tag{10}
\]
or
\[
y^A_1 > y^B_1, \quad y^A_1 \neq x^A_1, \quad \text{and} \quad y^B_1 \neq x^B_{j+1}. \tag{11}
\]
Since (10) and (11) are symmetric, we proceed by assuming (10). Then, by the feasibility of $y^A_1$ and $y^B_1$ in (1), we have $y^B_1 > x^B_1$ and $y^A_1 < x^A_{j+1}$. Since $(C, D)$ is the tightened pair of $(A, B)$, it follows from (3) that
\[
y^A_1 < \min(y^C_1, y^D_1) \leq \max(y^C_1, y^D_1) < y^B_1,
\]
which establishes (a). By the strict convexity of $L^j_1$, we obtain $L^j_1(y^A_1) + L^j_1(y^B_1) > L^j_1(y^C_1) + L^j_1(y^D_1)$. Therefore, we establish the strict inequality in (9) from the following result:
\[
L^j_1(y^A_1) + L^j_1(y^B_1) > L^j_1(y^C_1) + L^j_1(y^D_1) \quad \text{for each} \ j \text{ and } t \geq 2,
\]
7
which follows from $y_t^{A_j} + y_t^{B_j} = y_t^{C_j} + y_t^{D_j}$, Proposition 2 and the convexity of $L^j_t$. 

We immediately obtain the following result. In the statement of this result, we say a specific policy is suboptimal if there exists another policy that achieves a strictly lower expected cost in $[1, T]$ from at least one starting state.

**Theorem 4.** For the multi-echelon serial system described in Section 2, any policy outside the class of echelon order-up-to policies is suboptimal.

**Proof.** Inequality (9) implies either $\sum_{t=1}^{T^o} \phi_t(y_t^{A1}, y_t^{A2}) > \sum_{t=1}^{T^o} \phi_t(y_t^{C1}, y_t^{C2})$ or $\sum_{t=1}^{T^o} \phi_t(y_t^{B1}, y_t^{B2}) > \sum_{t=1}^{T^o} \phi_t(y_t^{D1}, y_t^{D2})$. The result now follows directly from the definition of a suboptimal policy. 

Our analysis is valid for discounted cost models for both finite and infinite horizons and also when linear ordering costs are present. Theorem 4 implies the optimality of echelon order-up-to policies when an optimal policy exists - this existence is guaranteed, for example, when demands are continuous random variables, by Corollary 8.5.2 and Proposition 9.17 of [1].

**Remark.** In this paper, we have shown the suboptimality of any policy that is not an echelon-order-up-to type when the single period cost function is strictly convex, under the setting of either the finite horizon or the infinite horizon discounted cost criterion. We explain why we are unable to prove the existence of an optimal echelon order-up-to policy when our assumptions are relaxed.

First, when the assumption of strict convexity of the single period cost is replaced with the assumption of weak convexity, our technique of working with the tightened pair leads to a new system which has a cost that is less than or equal to (as opposed to strictly less than) that of the original system. Moreover, since the number of possible states is infinite, there could be an infinite number of after-ordering inventory levels, and there is no guarantee that this process of working with tightened pairs will eventually converge to an echelon order-up-to policy. (Note that each step of the tightened pair “corrects” the optimal action at each pair of states, and a state-action pair that is tightened in one iteration may be subject to further tightening in a future iteration.) While it is plausible that such a difficulty can be circumvented by generalizing our technique to tighten infinitely many states or a continuum of states all at once, the complexity of such a technical argument is beyond the scope of this paper.
We now consider the long-run average cost criterion under the assumption that the single period cost function is strictly convex. While our analysis shows that, for any finite $T$, the $T$-period average cost of any system that does not follow an echelon order-up-to policy has a strictly higher cost than the tightened system (see (9)), it is not clear if the strict inequality holds for the infinite horizon long run average costs. For reasons similar to those discussed in the previous paragraph, we do not know if the iterative process of tightening will eventually result in an echelon order-up-to policy.

Acknowledgement

We sincerely thank the review team for its valuable suggestions that have improved the exposition of the paper. The first author was supported partially by NSF grant DMS-0732169.

References


