OPERATIONS RESEARCH
http://dx.doi.org/10.1287/opre.1120.1060ec

informs

e-companion

ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"On Optimal Policies for Inventory Systems with Batch Ordering" by Woonghee Tim Huh and Ganesh Janakiraman, *Operations Research*, http://dx.doi.org/10.1287/opre.1120.1060.

Online Appendices

A Detailed Comparison of Our Work with Chen (2000)

One of the crucial differences between our model and the model studied by Chen (2000) is, as noted in Section 2, that Chen's proof cannot be used when the assumption of IID demands in all periods is removed, which we explain further below. Our explanation involves a discussion on how Chen and Zheng (1994) compares with Clark and Scarf (1960) and a discussion on how the comparison between Chen's paper and our paper is similar. (For the sake of brevity, we will use **CS** to refer to Clark and Scarf (1960) and use **CZ** to refer to Chen and Zheng (1994).)

CS's proof of the result that echelon base-stock policies are optimal for serial inventory systems (without batch ordering constraints) has two main ideas. The first idea is based on the intuition that in any period, every echelon is constrained in its inventory decisions by the availability of inventory in its immediately upper echelon and that the upper echelon should be charged an appropriate *induced penalty cost function* for preventing the lower echelon from taking its optimal actions. This idea is important for the decomposition of the system's cost function into *convex* echelon-specific cost functions, about which we have mentioned in the main body of the paper. The second idea is that this decomposition is preserved from period to period in the dynamic program.

CZ re-derive **CS**'s result by assuming that demands in all periods are IID whereas **CS** only required independence. **CZ**'s proof relies heavily on the concept of *myopic optimality* (see Heyman and Sobel (1984)); this is the property that, if the cost in every period is a function (which is independent of the period index) of the state after the action in that period is taken and it is feasible in every period to reach the single period *cost-minimizing state*, then this myopic policy is an optimal policy for any planning horizon. The assumption of IID demands implies that the cost function in every period is the same, thus enabling **CZ** to use myopic optimality. Their proof then requires the study of a single period problem. They define this single period problem using a cost which is expressed in terms of echelon-N's inventory position alone (as opposed to the original decomposition into several cost functions, one for each echelon) by introducing a shift in the time indices associated with the original cost model (this shift in the time indices is inconsequential for the infinite horizon problem). Since the system's cost has now been expressed as a function of echelon-N's inventory position alone, and this function is the same in every period, it is feasible to raise echelon-N's inventory position to the minimizer of this function in every period, assuming the starting inventory position in the first period did not exceed this minimizer. Thus, this policy, which is an echelon base-stock policy, is optimal because the minimum value of the above cost function can be attained in every period.

To summarize the comparison of the proofs in **CS** and **CZ**, **CS** involves two ideas, (1) decomposition of the system's cost function into echelon-based functions and (2) the preservation of this decomposition by dynamic programming, whereas **CZ** only uses the first idea and then appeals to myopic optimality. In this sense, the latter proof with the assumption of IID demands is simpler than the original proof of **CS** (which only assumes independent demands).

The analysis of Chen (2000) is a generalization of **CZ**'s analysis to serial systems with batch ordering constraints. Again, the assumption of IID demands allows the use of the myopic optimality concept but in the following more general manner. Without the batch ordering constraint, it is feasible in every period to reach an after-ordering state which minimizes the single period cost which is a function of the echelon-N inventory position. The batch ordering constraint makes it necessary to consider not just this one state but an interval of states of length equal to Q^N, the batch size for echelon-N, such that the average cost incurred across this interval is the lowest. Chen's proof is based on demonstrating the following facts: (1) Under any feasible policy, for every i, the steady state distribution of the echelon-i inventory position (in *modular arithmetic with respect to* Q^i) is uniform over the set $\{0, 1, 2, \ldots, Q^i - 1\}$. (2) The average of the single period cost for the system (which has been expressed as a function of the echelon-N inventory position) over the integers $\{x, x+1, \ldots, x+Q^N-1\}$ is a quasi-convex function of x. Since this single period cost function is the same for every period, the policy of letting the inventory position reach its cost minimizing interval in every period is feasible and optimal. This establishes the optimality of an echelon (R, nQ) policy with Chen's assumption of IID demands.

In our work, on the other hand, we only assume that demands are independently distributed. Thus, the expected cost for the system as a function of the inventory positions in a period is not the same across time periods, implying that myopic optimality cannot be used. Moreover, even for the single stage problem (i.e. when $N = 1$), the cost-to-go function is not necessarily convex or quasi-convex. Thus, two key steps become necessary to provide a proof of the optimality of echelon (R, nQ) policies in our case. The first step is the identification of a property weaker than convexity such that an (R, nQ) policy can be shown to be optimal for a cost function that possesses this property. We identify this property to be the *Q-difference-increasing* property. The second step is to show that the **CS** analysis using dynamic programming (as opposed to the myopic optimality approach) and decomposition of cost functions continues to be useful with batch ordering. A crucial component of this step is to show that the echelon-based cost functions constructed in this approach have the Q-difference-increasing property *in every period.* To summarize the entire discussion above relating our work with **CS**, **CZ**, and Chen, our paper generalizes the work of **CS** to batch-ordering systems under independent demands while Chen generalizes the work of **CZ** to these systems under IID demands.

B Algorithmic Issues

In this section, we discuss how our results lead to efficient algorithms for computing the optimal policy for our system under both the finite horizon expected discounted cost criterion as well as the infinite horizon average cost per period criterion.

For the finite horizon problem, the algorithm of Clark and Scarf (1960) is *efficient* for serial systems without the batch ordering constraint in the following sense (we use *efficient* in the same sense throughout this section): While the original dynamic program has a state space of dimension N, the number of echelons, their algorithm involves solving N *one-dimensional* dynamic programs; thus, their algorithm avoids the *curse of dimensionality* which the original dynamic program suffers from. In our notation, for the case of integer demands, this algorithm computes the quantity r_t^i for every echelon i and period t with $Q^i = 1$ for each echelon i, and the resulting echelon order-up-to policy with these order-up-to levels is optimal. For systems with batch ordering, our results and analysis of Section 4 show that an analogous algorithm can be used to compute the quantity r_t^i for any echelon Q^i without increasing the computational effort and that the resulting echelon (R, nQ) policy, as described in Theorem 4.3, is optimal. In other words, we now know that the optimal policy for batch ordering systems can be computed by essentially using the efficient Clark-Scarf algorithm.

Next, we discuss the infinite horizon average cost case. While the Clark-Scarf algorithm is efficient for the finite horizon problem, its computational time increases linearly with the length of the horizon. Federgruen and Zipkin (1984) show that for the special case of IID demands, the Clark-Scarf algorithm is equivalent to solving a single period problem – this result holds for the infinite horizon average cost criterion, the infinite horizon discounted cost criterion and the finite horizon criterion (under a certain end-of-horizon assumption). The algorithm of Chen (2000) is a generalization of Federgruen and Zipkin's idea to *batch ordering* systems under *IID demands* and the infinite horizon average cost criterion. Chen's algorithm also reduces to solving a single period problem.

A paper closely related to this discussion is Chen and Song (2001). They study an infinite horizon average cost model *without batch ordering* in which demands are *not IID* – they model non-identical demands by a Markov-modulated process – there is a finite-state Markov chain whose state, w_t, is observed at the beginning of every period t, and the distribution of the random variable D_t depends only on w_t. They present an efficient algorithm for computing the base-stock levels for each echelon as a function of the state w in a period and show that the resulting echelon order-up-to policy is optimal. An intuitive description of their analysis follows.[3] The idea of decomposing the multi-echelon problem

[3]Note: Our discussion of Chen and Song's analysis here is only intended to provide the reader a sketch of the key ideas. A careful discussion would necessitate elaborate notation. The term "mega-period" is borrowed from Chen and Song's paper. We use the term "mega-period dynamic program" as a crude

into single echelon problems continues to be fundamental to the analysis. The *key new concept* they introduce is a *mega-period* and a corresponding cost minimization problem which we will refer to as the *mega-period dynamic program.* This mega period is the period of time between two consecutive visits of the Markov chain to a specific state, say $*^i$, for every echelon i. The state $*^i$ is defined in such a way that the order-up-to level of the first period of the mega-period dynamic program (by definition, the Markov state in this first period is $*^i$), say $r^i(*^i)$, is larger than the order-up-to levels of all other periods. Thus, under the optimal policy for this dynamic program, the echelon-i inventory is exactly $r^i(*^i)$ in every period t such that $w_t = *^i$ (i.e., beginning of a mega-period). This regeneration implies a generalization of the notion of *myopic optimality* which usually occurs in inventory models with IID demands, where the solution to the single period problem is optimal to the dynamic problem; in the Markov-modulated case, the optimal policy for the mega-period dynamic program turns out to be optimal to the infinite horizon average cost problem. The computational effort required by this algorithm is linear in N and is polynomial in the number of states of the modulating Markov chain.

Given this background, the immediate questions that arise are whether it is possible to use our results to obtain an algorithm for the infinite horizon average cost problem with batch ordering and Markov-modulated demand, and whether this algorithm is efficient in the sense that it does not substantially increase the computational complexity relative to Chen and Song's algorithm without batch ordering. The answers to these questions are "yes"; we explain how the results of this paper can be combined with the analysis of Chen and Song (2001) to produce such an algorithm. The concept of a mega-period (for each echelon) from Chen and Song's analysis is necessary here too. Solving every echelon's mega-period dynamic program is achieved using our results from Section 3. While Chen and Song show that the cost functions that appear in these dynamic programs are convex in every period, the corresponding cost functions with batch ordering are Q^i-difference-increasing, a fact which follows from our results in Section 4. Another important modification in the analysis with batch ordering is that, the echelon-i inventory at the beginning of every period t in which $w_t = *^i$ belongs to the interval $\{r^i(*^i), \ldots, r^i(*^i) + Q^i - 1\}$ – this necessitates the computation of certain cost functions related to the mega-period dynamic program at each of the Q^i integers in this interval – thus, the algorithm with batch ordering is slightly more involved than Chen and Song's algorithm – however, it is still efficient in that the computational effort increases in Q^N (the largest batch size) in a linear manner.

---

reference to an iterative cost accounting scheme they construct and certain optimization problems that appear in that scheme.

C Proofs

Proof of Lemma 3.1

Proof. We first show statement (i). Let $y = s(x) > -\infty$. It is sufficient to show the claims that (a) $g(y) \leq g(y+Q)$ and (b) for all $y' \leq y$ such that $[y']^Q = x$, we have $g(y') - g(y'-Q) < 0$. To show (a), notice that $y \geq r$; thus, by the Q-difference-increasing property of g, we know that $g(y+Q) - g(y) \geq g(r+Q) - g(r) \geq 0$, where the last inequality follows from the definition of r. We now proceed to (b). For all $y' \leq y$, $g(y') - g(y'-Q) \leq g(y) - g(y-Q)$ by the Q-difference-increasing property of g. Thus, it is enough to show that $g(y) - g(y-Q) < 0$. Define $\tilde{y} = y - Q$ and note that $\tilde{y} \leq r$. Thus, by the definition of r, we observe that $g(y) = g(\tilde{y} + Q) < g(\tilde{y}) = g(y - Q)$ and prove (b). This completes the proof of the first statement in the lemma.

Statement (i) implies that, for any $x \in [0, Q)$, the function $g(y)$ defined over the set of points $\{y : [y]^Q = x\}$ is decreasing until $s(x)$ and is increasing thereafter. Statements (ii) and (iii) are direct consequences of this property. □

Proof of Proposition 3.2

Proof. Since $f^1_{T+1}(\cdot) = 0$, it is trivially Q^1-difference-increasing. We proceed inductively by assuming the statements in the proposition hold for $t+1$ and establishing them to be true for t. By this inductive hypothesis, we know that $f^1_{t+1}(\cdot)$ is Q^1-difference-increasing. It is straightforward to verify that the Q^1-difference-increasing property is preserved under the expectation operator and the addition operator, and thus it follows from (3.2) and the Q^i-difference-increasing property of G^1_t that $g^1_t(\cdot)$ is also Q^1-difference-increasing.

Next, we define $s^1_t([x^1_t]^{Q^1})$ as the unique number s such that $r^1_t \leq s < r^1_t + Q^1$ and $[s]^{Q^1} = [x^1_t]^{Q^1}$. Then, statement (ii) of Lemma 3.1 implies that the solution to the optimization problem in (3.1) is given by

$$y^1_t \quad = \quad \begin{cases} x^1_t & \text{if } x^1_t > r^1_t \\ s^1_t([x^1_t]^{Q^1}) & \text{if } x^1_t \leq r^1_t \end{cases} ,$$

This establishes (ii) for period t.

It only remains to show that $f^1_t(\cdot)$ is Q^1-difference-increasing. Substituting the above solution for y^1_t into (3.1), we get

$$f^1_t(x^1_t) \quad = \quad g^1_t(\max\{s^1_t([x^1_t]^{Q^1}), x^1_t\}) \ .$$

Then, for any $x^1_t < r^1_t$, it follows from the definition of $s^1_t([x^1_t]^{Q^1})$ that $x^1_t < x^1_t + Q^1 \leq s^1_t([x^1_t]^{Q^1})$, and thus,

$$f^1_t(x^1_t + Q^1) - f^1_t(x^1_t) \quad = \quad g^1_t(s^1_t([x^1_t]^{Q^1})) - g^1_t(s^1_t([x^1_t]^{Q^1})) \ = \ 0 \ .$$

For $x_t^1 \geq r_t^1$, we have $s_t^1([x_t^1]^{Q^1}) \leq x_t^1 < x_t^1 + Q^1$ and thus obtain

$$f_t^1(x_t^1 + Q^1) - f_t^1(x_t^1) \;=\; g_t^1(x_t^1 + Q^1) - g_t^1(x_t^1)\;.$$

This expression is increasing in x_t^1 by the Q^1-difference-increasing property of g_t^1, and non-negative by the definition of r_t^1. Thus, we conclude that $f_t^1(\cdot)$ is Q^1-difference-increasing, as required. This completes the induction argument and the proof of the proposition. □

Proof of Proposition 4.1

Proof. We need to prove that, for any $x' \leq x''$,

$$\lambda_t^2(x' + Q^1) - \lambda_t^2(x') \;\leq\; \lambda_t^2(x'' + Q^1) - \lambda_t^2(x'')\;. \tag{C.1}$$

Using the definition of λ_t^2 and the fact that $[x+Q]^Q = [x]^Q$, we can verify that this inequality is equivalent to the following:

$$\begin{aligned}
&\min_y \;\left\{g_t^1(y) \mid y \leq x' + Q^1,\; [y]^{Q^1} = [x']^{Q^1}\right\} - \min_y \;\left\{g_t^1(y) \mid y \leq x',\; [y]^{Q^1} = [x']^{Q^1}\right\} \\
&\quad\leq\; \min_y \;\left\{g_t^1(y) \mid y \leq x'' + Q^1,\; [y]^{Q^1} = [x'']^{Q^1}\right\} - \min_y \;\left\{g_t^1(y) \mid y \leq x'',\; [y]^{Q^1} = [x'']^{Q^1}\right\}\;.
\end{aligned}$$

Using statement (iii) of Lemma 3.1, this inequality can be rewritten as

$$\begin{aligned}
&g_t^1(\min\{x' + Q^1, s_t^1([x']^{Q^1})\}) - g_t^1(\min\{x', s_t^1([x']^{Q^1})\}) \\
&\quad\leq\; g_t^1(\min\{x'' + Q^1, s_t^1([x'']^{Q^1})\}) - g_t^1(\min\{x'', s_t^1([x'']^{Q^1})\}).
\end{aligned} \tag{C.2}$$

We now consider three exhaustive cases:

- Case $r_t^1 \leq x' \leq x''$: In this case, $x' \geq s_t^1[x']$ and $x'' \geq s_t^1[x'']$. Now, the left and right hand sides of (C.2) become zero and the desired inequality holds.
- Case $x' < x'' < r_t^1$: In this case, $x' + Q^1 \leq s_t^1[x']$ and $x'' + Q^1 \leq s_t^1[x'']$. Now, (C.2) is equivalent to
  $$g_t^1(x' + Q^1) - g_t^1(x') \leq g_t^1(x'' + Q^1) - g(x'')\;,$$
  which is true because of the Q^1-difference-increasing property of g^1.
- Case $x' < r_t^1 \leq x''$: In this case, the left hand side of (C.2) is $g_t^1(x'+Q^1) - g_t^1(x')$ which is negative because $x' < r_t^1$ (this fact is easily seen from the definition of r_t^1). The right hand side is zero. Thus, the desired inequality holds.

Therefore, we obtain (C.2), and we complete the proof. □

Proof of Proposition 4.2

Proof. Since $\lambda_t^2(\cdot)$ is Q^1-difference-increasing (Proposition 4.1) and Q^2 is a positive integer multiple of Q^1, it follows that $\lambda_t^2(\cdot)$ is also Q^2-difference-increasing. The Q^2-difference-increasing property of $f_t^2(\cdot)$, as well as the structure of the optimal policy, can be shown inductively in the same way as in the proof of Proposition 3.2. $\square$

Proof of Theorem 4.3

Proof. Statements (i) and (ii) hold trivially for $t = T$ because f_{T+1}° is the zero function and G_T^i is a Q^i-difference-increasing function for $t \in \{1, 2\}$. Statement (iii) is easy to verify for $t = T$ from the definitions of f_T°, f_T^1, f_T^2 and λ_T^2. For $t \in \{1, \ldots, T-1\}$, we assume (i), (ii) and (iii) hold for $t+1$, and proceed by induction to prove these statements for t. Recall that the before-ordering echelon inventory vector, denoted by (x_t^1, x_t^2), satisfies $[x_t^2]^{Q^1} = [x_t^1]^{Q^1}$ (Remark 2). From the definition of $g_t^\circ(y_t^1, y_t^2)$ in (4.5) and the induction hypothesis (iii) for f_{t+1}°, $g_t^\circ(y_t^1, y_t^2)$ can be written as

$$\begin{aligned}
&G_t^1(y_t^1) \ + \ \alpha E\left[f_{t+1}^1(y_t^1 - D_t)\right] \\
&\qquad + \ G_t^2(y_t^2) \ + \ \alpha E\left[\lambda_{t+1}^2(y_t^2 - D_t) + f_{t+1}^2(y_t^2 - D_t)\right] \ .
\end{aligned}$$

Note that the sum of the first two terms coincides with the definition of $g_t^1(x_t^1)$ in (3.2); similarly, the third and the fourth terms sum up to $g_t^2(x_t^2)$ given in (4.3). Thus, we prove (i) for t.

Therefore, (4.6) can be written as

$$\begin{aligned}
&f_t^\circ(x_t^1, x_t^2) \\
&\quad = \min_{y_t^1, y_t^2} \quad g_t^1(y_t^1) + g_t^2(y_t^2) - c_t^1 \cdot x_t^1 - c_t^2 \cdot x_t^2 \\
&\qquad\quad \text{s. t.} \quad x_t^1 \le y_t^1 \le x_t^2 \le y_t^2, \quad [y_t^1]^{Q^1} = [x_t^1]^{Q^1}, \quad [y_t^2]^{Q^2} = [x_t^2]^{Q^2} \ .
\end{aligned}$$

In the optimization problem above, the objective function is separable, and the constraints can also be decoupled. Thus, the optimal solution can be found by solving the following two single-variable optimization problems:

$$\min_{y_t^2} \ \left\{ g_t^2(y_t^2) \mid y_t^2 \ge x_t^2, \ [y_t^2]^{Q^2} = [x_t^2]^{Q^2} \right\} - c_t^2 \cdot x_t^2 \ , \qquad \text{and} \tag{C.3}$$

$$\min_{y_t^1} \ \left\{ g_t^1(y_t^1) \mid x_t^2 \ge y_t^1 \ge x_t^1, \ [y_t^1]^{Q^1} = [x_t^1]^{Q^1} \right\} - c_t^1 \cdot x_t^1 \ . \tag{C.4}$$

Moreover, $f_t^\circ(x_t^1, x_t^2)$ is the sum of the expressions given in (C.3) and (C.4). The first problem (C.3) is the same as $f_t^2(x_t^2)$ given in (4.2). Thus, the optimal value of y_t^2 is the same as the solution given in Proposition 4.2. The second problem (C.4) is the same as $f_t^1(x_t^1)$ in (3.1) except for the upper bound constraint $y_t^1 \le x_t^2$. The fact that g_t^1 is Q^1-difference-increasing (by Proposition 3.2) and statement (iii) of Lemma 3.1 together imply the following: the optimal policy for y_t^1 is to order the minimum number of batches such that y_t^1 is at least r_t^1, if attainable, and otherwise to order as much as possible (up to the upper bound x_t^2). This proves (ii) for t.

From the above discussion, $f_t^\circ(x_t^1, x_t^2)$ is the sum of optimal values for (C.3) and (C.4). Note that the optimal value for (C.3) is $f_t^2(x_t^2)$ by (4.2). Also, (C.4) can be written as

$$f_t^1(x_t^1) \ + \ \lambda_t^2(x_t^2) \ ,$$

from the definition of $f_t^1(x_t^1)$ in (3.1) and the definition of $\lambda_t^2(x_t^2)$ in (4.1). Thus,

$$f_t^\circ(x_t^1, x_t^2) \ \ = \ \ f_t^2(x_t^2) \ + \ f_t^1(x_t^1) \ + \ \lambda_t^2(x_t^2) \ ,$$

and we complete the proof of (iii). □