



Manufacturing & Service Operations Management

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Utilizing Forecast Band Refinement for Capacitated Production Planning

Philip Kaminsky, Jayashankar M. Swaminathan,

To cite this article:

Philip Kaminsky, Jayashankar M. Swaminathan, (2001) Utilizing Forecast Band Refinement for Capacitated Production Planning. Manufacturing & Service Operations Management 3(1):68-81. <http://dx.doi.org/10.1287/msom.3.1.68.9999>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

© 2001 INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Utilizing Forecast Band Refinement for Capacitated Production Planning

Philip Kaminsky • Jayashankar M. Swaminathan

Department of IEOR, University of California, Berkeley, California 94720

The Kenan-Flagler Business School, University of North Carolina, Chapel Hill, North Carolina 27599

kaminsky@ieor.berkeley.edu • msj@haas.berkeley.edu

We present a model for forecast evolution that captures two notions related to forecasts: (1) forecasts are not exact; (2) forecasts over longer horizons are less certain than those over shorter horizons. We model the forecast of discrete demand as a band defined by the lower and upper bounds on demand, such that future forecasts lie within the current band. We develop a capacitated production planning model for a single product with terminal demand. We develop four heuristics for the problem and characterize their performance. In particular, two of the heuristics are optimal for the no holding-cost case. In our computational study, we analyze the performance of our heuristics and compare them to the optimal solution and to a simple heuristic that simulates common industrial practice using point forecasts. We find that two of our heuristics are very close to the optimal solution (less than 0.5% away from optimal on average under the conditions studied). Further, we consider forecast update patterns with primarily early, intermediate, and late information updates and provide insights on the effect of information update patterns on optimal costs. (*Forecast Evolution; Production Planning; Discrete Demand; Capacity; Heuristics; Information Updates*)

1. Introduction

In the past few years, there has been increasing emphasis placed on improving the quality of forecasts within a supply chain (Lee et al. 1997). It is becoming more important to develop models and techniques that enable firms to make effective use of these forecasts in production planning. Any approach should account for the following two facts about forecasts, often stated colloquially as follows (for example, Simchi-Levi et al. 2000): (1) the forecast is always wrong; (2) the longer the forecast horizon is, the more wrong the forecast is. Traditionally, a "forecast" consists of a single number, or point estimate, that is revised over time. Our recent conversations with the manufacturing manager of a major electronics company have highlighted this fact. The manager pointed out to us that although the forecasts are continuously revised, the firm has found that the second to last es-

timate is often the most accurate! Their forecasts often show swings from period to period (this is particularly true for high-tech products) due to changes in market conditions, developments in technology, and changes in competitors pricing policies, among other reasons. As firms realize that Point (1) (above) is true, many of them are moving toward a concept similar to a forecast band, with optimistic and pessimistic forecasts or, perhaps, a 3-point distribution with optimistic, pessimistic, and most likely values. These forecast bands are updated each period as new information arrives. The changing bands provide manufacturing with more information than in the past; it is up to the manufacturer to make efficient and effective production decisions using this information.

In this paper, we present a model that focuses on capturing Point (2) (above), that as forecast horizons become longer, forecasts become more uncertain. In

a capacitated environment, there is a trade-off between capacity and good forecast information, as there may not be time to wait for the forecast to get "good enough." In particular we consider a forecast evolution model that is defined by a band and a width that captures the uncertainty in forecast. As time moves forward, subsequent forecasts have smaller widths (representing a better forecast) and are contained within the band defined by earlier forecasts. This models a forecasting process that gets refined over time as new information arrives. The manufacturing firm utilizing this forecast has a fixed capacity in each period and needs to decide in which periods it should produce, taking into account expected production, holding, salvage, and stock-out costs. In this paper, we model products for which demand occurs at the end of the horizon and updated forecast information is obtained each period. We prove the existence of inventory threshold levels in each period, below which the firm should produce and above which it should not. We provide four simple heuristics and show their relationship to the optimal production decision. In particular, two of the heuristics are equivalent and optimal in the zero-holding cost case. In our computational study, we compare the performance of a practical heuristic that considers only a point forecast estimate with heuristics that consider the additional information in the bands under different conditions related to costs, time horizon, timing of forecast updates, and capacity restrictions. Further, we compare the performance of our heuristics with the optimal solution (obtained by a dynamic programming recursion). Finally, we develop an algorithm for the multiunit capacity case with zero-holding cost and demonstrate its optimality.

The paper is organized as follows: In §2 we present related literature. In §3 we present the basic model, results, and heuristics. In §4, we present our computational results and insights. In §5, we consider the multiple capacity case and extend our results. We conclude in §6.

2. Related Literature

Forecast evolution models have been studied by several researchers in the past. Hausman (1969) studies

the problem in which improved forecasts are available before each decision stage, and he models the evolution of forecasts as a quasi-Markovian or Markovian system. He suggests modeling a series of ratios of successive forecasts as independent lognormal variates and presents a dynamic programming formulation. Hausman and Peterson (1972) consider a capacity constrained multiproduct system with terminal demand with the forecasts for total sales following a lognormal model. They formulate the problem as a dynamic program and provide heuristics for solving the problem. Heath and Jackson (1994) introduce a martingale model for forecast evolution (MMFE) and present simulation results. Gullu (1996) considers a special case of the MMFE model and shows that in a capacitated environment the system performs better when one-period demand forecast is utilized than when it is not. More recently, Toktay and Wein (1999) consider an MMFE model and characterize effective policies under heavy traffic assumptions for a capacitated single server.

Bayesian models for forecast updates in a inventory setting were first studied by Scarf (1959). Azoury (1985) extends some of Scarf's results. Bitran et al. (1986) study a terminal demand model for style goods under capacity restrictions and Bayesian forecast updates, where each of the product families is produced one period before the realization of demand. They present a stochastic mixed integer programming formulation and provide a decomposition scheme and bounds on the optimal solution. Fisher and Raman (1996) consider a two-period problem from the fashion goods industry where the manufacturer needs to determine production quantities in two periods, with lower production costs in the first period but with improved forecast information available in the second. Agarwal et al. (1999) consider a similar problem but analyze capacity outsourcing issues. Nurani et al. (1994), Lee and Whang (1998), and Gurnani and Tang (1999) have investigated uncapacitated inventory problems with forecast updates. In addition, capacitated nonstationary inventory problems have been studied by several authors, including the Federgruen and Zipkin (1986), Morton and Pentico (1995), Aviv and Federgruen (1997), Kapuscinski and Tayur (1998), and Gavirneni et al. (1999).

In this paper, we introduce a forecast evolution process based on bands, in which the range of possible demand is completely contained within a *forecast band*. This band, defined by upper and lower bounds, captures some of the important elements of real-world forecast revisions and, in particular, the concept of improved information over time (Buckley et al. 1995). It is to be noted that earlier papers such as Bitran et al. (1986) and Hausman and Peterson (1972) consider forecast evolutions that get updated and refined over time. The former captures forecast evolution through standard deviation from the actual demand, while the latter captures the evolution through the relative uncertainty in the forecast. Our model differs in that we restrict the future demand forecasts using a hard limit on the lower and upper bounds that allows us to simplify the capacitated production planning problem. In a multiproduct environment, Bitran et al. (1986) consider set up and holding costs, whereas Hausman and Peterson (1972) consider only the production costs. Both papers focus on developing effective heuristics for the problem. In contrast to these papers, we consider a single product environment and provide an optimal solution to the no holding-cost problem without the explicit necessity for a dynamic program. We also develop effective heuristics based on that for the case with holding costs.

3. Basic Model

3.1. Forecast Evolution

In this section we present the forecast evolution model with discrete demand. We number time periods in reverse order, so that when we say time t , we mean t periods remain. Our notation is summarized below:

- a_t : lower bound on the demand forecast with t periods to go (a_0 represents realized demand);
- w_t : width of the forecast with t periods to go (w_0 equal to zero);
- $F_t(a_t)$: probability distribution of a_{t-1} in the interval $[a_t, a_t + w_t - w_{t-1}]$;
- α_t : known reduction in width of forecast with t periods to go;
- x_t : on-hand inventory with t periods to go;

- π : per unit penalty cost for demand not met;
- c : per unit cost of production;
- h : per unit holding cost incurred on inventory held at the end of any period;
- $s(< c)$: per unit salvage cost for each unit remaining after demand is met;
- $V_t(x_t, a_t, w_t)$: expected optimal cost incurred with t periods to go.

In the forecast evolution, we seek to capture two important characteristics: (1) the forecast gets tighter as one gets closer to demand; (2) the future forecasts are contained in the current range. We assume that given that we are at time t , the final demand can only take on discrete values in the interval $[a_t, a_t + w_t]$. Each period the range of possible demand gets smaller as $t \rightarrow 0$. That is, $w_{t-1} = w_t - \alpha_t$, where each $\alpha_t \geq 0$, $\forall t$ are such that $w_1 \geq 0$. Note that we assume that the progression of w_t is known, whereas the progression of a_t is probabilistic. This models a situation in which the firm has a good understanding of when additional information will be obtained (for example, a trade show) but doesn't know *a priori* what that information will be. Thus, given the range of possible demands at time t , the range of possible demands at time $t - 1$ is $w_t - \alpha_t$ wide and is contained within the range $[a_t, a_t + w_t]$. α_t provides an ability to model the different points of time at which major or minor updates in forecast may be obtained. For example, a firm that obtains all the forecast updates very close to the selling season will have $\alpha_t = 0$ for higher values of t , and a high value for α_t for lower values of t . The final demand a_0 is defined on the discrete interval $[a_1, a_1 + w_1]$ with a probability distribution $F_1(a_1)$. $F_t(a_t)$, known *a priori*, provide us with an ability to model changes in distribution over time.

3.2. Production Model

In our model we assume that demand occurs at the end of the horizon, but periodic forecast updates are obtained about the forecast in each period. This situation is very common in several industries in which the product selling season has a short time frame. To simplify our analysis, we assume that the manufacturer has a unit production capacity each period and that the final demand is in multiples of this capacity.

We extend this model to the case where the firm has multiple units of capacity in §5. Note that if there were no capacity restriction, one would wait until the final period and then produce the single period newsboy amount. The total cost at time $t = 0$, when a demand a_0 is realized, is given by $V_0(x_0, a_0, 0) = s(x_0 - a_0)^+ + \pi(a_0 - x_0)^+$. To simplify the notation, we do not explicitly represent the distribution $F_t(a_t)$ in our value function. In our subsequent computational analysis, we assume $F_t(a_t)$ to be uniformly distributed for all t .

$$V_t(x_t, a_t, w_t) = \min \begin{cases} c + h(x_t + 1) \\ + E_{a_{t-1}}\{V_{t-1}(x_t + 1, a_{t-1}, w_t - \alpha_t)\} \\ h(x_t) + E_{a_{t-1}}\{V_{t-1}(x_t, a_{t-1}, w_t - \alpha_t)\}. \end{cases}$$

Let $\tilde{V}_0(x_0, a_1, w_1)$ represent the single period cost, given that the demand is distributed between a_1 and $a_1 + w_1$, p_i is the probability that demand is equal to i , and the on-hand inventory is x_0 . The single period cost is given by

$$\begin{aligned} \tilde{V}_0(x_0, a_0, w_1) &= E_{a_0}\{V_0(x_0, a_0, 0)\} \\ &= \sum_{a_0=a_1}^{a_1+w_1} p_{a_0}\{s(x_0 - a_0)^+ + \pi(a_0 - x_0)^+\}. \end{aligned}$$

The following results characterize the structure of the optimal policy for this model.¹

LEMMA 3.1. $E_{a_0}\{V_0(x + 1, a_0, 0) - V_0(x, a_0, 0)\}$ is monotone increasing in x , i.e., $E_{a_0}\{V_0(x + 2, a_0, 0) - V_0(x + 1, a_0, 0)\} \geq E_{a_0}\{V_0(x + 1, a_0, 0) - V_0(x, a_0, 0)\} \forall x$.

Note that Lemma 3.1 is analogous to the cost function being convex for the continuous case.

THEOREM 3.2. For every period t there exists a threshold level $y_t^*(a_t, w_t)$, such that it is optimal to produce in period t if $x_t < y_t^*(a_t, w_t)$ and to not produce if $x_t \geq y_t^*(a_t, w_t)$.

We have shown that the optimal production policy has a threshold form in that it is optimal to produce in a period if $x_t < y_t^*$ and not to produce if $x_t \geq y_t^*$. As with any dynamic program, computing the opti-

mal policy can be quite time and memory intensive. In particular, for the dynamic program to solve quickly it is necessary to store $O(n^2 a_n)$ values. (Furthermore, the multiple capacity case considered in §5 will be even more memory intensive.) This motivated the development of several heuristics for the problem that we present in the following subsection.

3.3. Heuristics

3.3.1. Heuristics HUB and HLB. In this section, we present two simple and easily implementable heuristics for deciding on the production rule for any period. These heuristics, rather than considering all future time periods, base the threshold level for the current period on the probability distribution of the forecast in period 1 given the probability distribution in the current period n , i.e., the decision to produce only depends on a_n (and w_t $0 \leq t \leq n$) and not on intermediate values. This makes these heuristics much faster computationally than a dynamic program. Somewhat surprisingly, this heuristic is very effective, particularly in the case when holding cost equals zero.

Step 1. For $a_1 \in \{a_n, a_n + 1, \dots, (a_n + w_n - w_1)\}$.

Step 1a. Calculate x^* , taking current inventory of x_n into account. Holding cost is accounted for as if the quantity desired is made as late as possible. Specifically,

$$x^* = \operatorname{argmin}_{a_1 \leq y \leq a_1 + w_1} \left\{ \tilde{V}_0(y, a_1, w_1) + c(y - x_n)^+ + h \frac{(y - x_n)^+(y - x_n + 1)^+}{2} \right\},$$

where $\tilde{V}_0(x, a_1, w_1)$ is the single period remaining cost if inventory is x , and demand can take on the discrete values between a_1 and $a_1 + w_1$ with a probability distribution given by $F_1(a_1)$.

Step 1b. The following three cases are considered:

Step 1b.1. If $x_n \geq x^*$, then

$$\begin{aligned} \lambda_{a_1} &= c + nh + \tilde{V}_0(x_n + 1, a_1, w_1) \\ &\quad - \tilde{V}_0(x_n, a_1, w_1). \end{aligned}$$

Step 1b.2. If $x_n + n \leq x^*$, then

¹The proofs of this and many of the subsequent lemmas, properties, and theorems are located in the appendix.

$$\lambda_{a_1} = c + nh + \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1).$$

Step 1b.3. If $x_n < x^* < x_n + n$, one of the following two penalties is charged:

(i) HLB:

$$\underline{\lambda}_{a_1} = nh + \frac{(x^* - x_n - 1)(x^* - x_n)}{2}h - \left\{ \frac{(x^* - x_n)(x^* - x_n + 1)}{2}h + (n - (x^* - x_n) - 1)(x^* - x_n)h \right\},$$

(ii) HUB:

$$\bar{\lambda}_{a_1} = \left\{ \frac{(x^* - x_n)(x^* - x_n + 1)}{2}h + (n - (x^* - x_n))(x^* - x_n)h \right\} - \left(\frac{(x^* - x_n)(x^* - x_n + 1)}{2}h \right).$$

Step 2. Let $p_{a_1|a_n}$ be the conditional probabilities of each a_1 given the current a_n . Let $\lambda_n = \sum_{a_1=a_n}^{a_n+w_n-w_1} p_{a_1|a_n} \lambda_{a_1}$ be the weighted average value of λ_{a_1} .

Step 3. If $\lambda_n \geq 0$, then don't produce in this period, otherwise produce.

The algorithms first compute the probability (denoted by $p_{a_1|a_n}$) for the lower bound in the last period a_1 given the current bound a_n using the F_t for all t from n to 1 and the values of a_t and w_t ($t \leq n$) in Step 1. Computationally, the probability values are easy to calculate because all values are discrete. Step 1a in the heuristic corresponds to determining the optimal inventory on hand in the final period for a given a_1 and accounting for holding cost in previous periods. In Step 1b an approximate marginal value for producing an extra unit in the current period is computed for three cases. For Case 1b.1, knowing that for the given a_1 , this additional unit is not going to be useful, the difference presents a penalty of $c + nh + \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)$ for producing that unit. For Case 1b.2, because if one unit is produced

every period for the next n periods, the final inventory will still be less than optimal, the heuristic assigns a penalty of $c + nh + \tilde{V}_0(x_n + n) - \tilde{V}_0(x_n + n - 1)$. For Case 1b.3, the calculation of this marginal value is not trivial because the costs depend on which subset of periods the actual production took place. As a result, we develop two values: a lower bound in which production occurs as late as possible and an upper bound in which production occurs as early as possible. If Step 1b.3.i is selected, we call the heuristic HLB, and if Step 1b.3.ii is selected, we call the heuristic HUB. Roughly speaking, Step 1b.3.i has two components. The first component

$$nh + \frac{(x^* - x_n - 1)(x^* - x_n)}{2}h,$$

represents an estimate of the cost associated with making one item now and the remaining $x^* - x_n - 1$ items as late as possible. The second component

$$\frac{(x^* - x_n)(x^* - x_n + 1)}{2}h + (n - (x^* - x_n) - 1)(x^* - x_n)h,$$

represents an approximation of the cost associated with making all $x^* - x_n$ items starting one period from now, in period $n - 1$. Similarly, Step 1b.3.ii has two components. The first component

$$\frac{(x^* - x_n)(x^* - x_n + 1)}{2}h + (n - (x^* - x_n))(x^* - x_n)h,$$

represents an approximation of the cost of making all $x^* - x_n$ items as soon as possible. The second component,

$$\frac{(x^* - x_n)(x^* - x_n + 1)}{2}h,$$

represents an approximate cost for making all $x^* - x_n$ items as late as possible. In Step 2 the expected value of the marginal difference is computed, and a decision to produce is made based on that. For future reference, let λ_n be represented by $\bar{\lambda}_n$ if 1b.3.ii is utilized and by $\underline{\lambda}_n$ if 1b.3.i is selected. Note that if the holding cost is zero, HLB and HUB are the same. For this zero-holding cost case, we prove the following:

THEOREM 3.3. *This heuristic is optimal when the holding cost equals zero.*

To prove this theorem, we need the following series of results.

PROPERTY 3.4. *When the holding cost equals zero, (i) $\lambda_n \geq E_{a_1|a_n}\{c + \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)\}$; (ii) $\lambda_n \leq E_{a_1|a_n}\{c + \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1)\}$.*

PROPERTY 3.5. *When the holding cost equals zero, (i) If $E_{a_{n-k}|a_n}\{c + V_{n-k}(x + k, a_{n-k}, w_{n-k}) - V_{n-k}(x + k - 1, a_{n-k}, w_{n-k})\} \geq 0$, then $E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + k - 1, a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x + k - 2, a_{n-k+1}, w_{n-k+1})\} \geq 0$; (ii) if $E_{a_{n-k}|a_n}\{c + V_{n-k}(x + 1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})\} \leq 0$, then $E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + 1, a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \leq 0$.*

Repeated application of Property 3.5 leads to the following lemma, which relates the expected marginal cost of producing an extra unit in the current period conditioned on a_1 , the final value of a , with the same cost conditioned on the value of a in the next period, a_{n-1} .

LEMMA 3.6. *When holding cost equals zero, (i) if $E_{a_1|a_n}\{c + \tilde{V}_0(x + n, a_1, w_1) - \tilde{V}_0(x + n - 1, a_1, w_1)\} \geq 0$, then $E_{a_{n-1}|a_n}\{c + V_{n-1}(x + 1, a_{n-1}, w_{n-1}) - V_{n-1}(x, a_{n-1}, w_{n-1})\} \geq 0$; (ii) if $E_{a_1|a_n}\{c + \tilde{V}_0(x + 1, a_1, w_1) - \tilde{V}_0(x, a_1, w_1)\} \leq 0$, then $E_{a_{n-1}|a_n}\{c + V_{n-1}(x + 1, a_{n-1}, w_{n-1}) - V_{n-1}(x, a_{n-1}, w_{n-1})\} \leq 0$.*

Combining Lemma 3.6 and Property 3.4 concludes the proof of Theorem 3.3. \square

The following result partially characterizes the effectiveness of the heuristic in the more general case with nonzero holding costs:

THEOREM 3.7. *If $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)\} \geq 0$, then HUB is guaranteed to provide an optimal decision, and if $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1)\} \leq 0$, then HLB is guaranteed to provide an optimal decision.*

To prove this theorem, we use the following series of results.

PROPERTY 3.8. *(i) If x_n represents the inventory with n periods to go, then $\lambda_n \geq E_{a_1|a_n}\{c + nh + \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)\}$; (ii) if x_n represents the inventory*

with n periods to go, then $\lambda_n \leq E_{a_1|a_n}\{c + nh + \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1)\}$.

PROPERTY 3.9. *(i) If $E_{a_{n-k}|a_n}\{c + kh + V_{n-k}(x + 1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})\} \geq 0$, then $E_{a_{n-k+1}|a_n}\{c + (k - 1)h + V_{n-k+1}(x + 1, a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \geq 0$; (ii) if $E_{a_{n-k}|a_n}\{c + kh + V_{n-k}(x + k, a_{n-k}, w_{n-k}) - V_{n-k}(x + k - 1, a_{n-k}, w_{n-k})\} \leq 0$, then $E_{a_{n-k+1}|a_n}\{c + (k - 1)h + V_{n-k+1}(x + k - 1, a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x + k - 2, a_{n-k+1}, w_{n-k+1})\} \leq 0$.*

LEMMA 3.10. *(i) If $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + 1, a_1, w_1) - \tilde{V}_0(x, a_1, w_1)\} \geq 0$, then $E_{a_{n-1}|a_n}\{c + h + V_{n-1}(x + 1) - V_{n-1}(x)\} \geq 0$; (ii) if $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + n, a_1, w_1) - \tilde{V}_0(x + n - 1, a_1, w_1)\} \leq 0$, then $E_{a_{n-1}|a_n}\{c + h + V_{n-1}(x + 1) - V_{n-1}(x)\} \leq 0$.*

Clearly, Lemma 3.10 results from repeatedly applying Property 3.9. Combining Lemma 3.10 and Property 3.8 proves Theorem 3.7. \square

3.4. Heuristics HCU and HCL

The preceding analysis suggests another set of heuristics. Recall that by Lemma 3.10, if $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + 1, a_1, w_1) - \tilde{V}_0(x, a_1, w_1)\} \geq 0$, then it is optimal to not produce, and if $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + n, a_1, w_1) - \tilde{V}_0(x + n - 1, a_1, w_1)\} \leq 0$, then it is optimal to produce. This suggests the following set of heuristics:

Step 1: Determine $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + 1, a_1, w_1) - \tilde{V}_0(x, a_1, w_1)\}$.

Step 2: Determine $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + n, a_1, w_1) - \tilde{V}_0(x + n - 1, a_1, w_1)\}$.

Step 3: If $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + 1, a_1, w_1) - \tilde{V}_0(x, a_1, w_1)\} \geq 0$, then don't produce. Or, if $E_{a_1|a_n}\{c + nh + \tilde{V}_0(x + n, a_1, w_1) - \tilde{V}_0(x + n - 1, a_1, w_1)\} \leq 0$, then produce. Or, either:

Step 3. HCU: Produce.

Step 3. HCL: Don't produce.

If Step 3 HCU is selected, we call this heuristic HCU. If Step 3 HCL is selected, we call this heuristic HCL. Note that if Step 3 HCU/HCL is not reached, this heuristic makes the optimal decision by Lemma 3.10. If they are reached, HCU will (obviously) tend to overproduce, while HCL will tend to underproduce.

4. Computational Study

In this section, we test our heuristics under different settings and explore the benefits of improving forecasts earlier in the time horizon. We use the following parameters for our computational study.

Time Horizon. We set the number of time periods T equal to 4 for SHORT horizons, 8 for MEDIUM horizons, and 12 for LONG horizons. We also used longer horizons up to $T = 20$ in experiments for analyzing the value of receiving earlier information updates.

Cost Parameters. We use production cost $c = 50$, penalty $\pi = 75, 150, 250$, salvage $s = 0, 10, 25, 45$, and holding cost $h = 0, 2, 4, 8, 12$ in the initial study (similar to Morton and Pentico 1995).

Information Update. We model the information updates in the forecast process by changing the width of the forecast band from period to period. We consider EARLY INFORMATION, INTERMEDIATE INFORMATION, and LATE INFORMATION models. In the EARLY (INTERMEDIATE, LATE) INFORMATION case, the band width reduces more per period in the first few (middle, last few) periods, respectively. For example, with MEDIUM horizon problems (8 periods and initial width of 11), the EARLY INFORMATION is represented by 2211100, which means that in the first 2 periods there is a reduction of 2 (that is the width decreases by 2 between each pair of time periods), and then in the next 3 periods there is a reduction of 1 each, and in the final 2 periods no additional information is obtained. On the other hand, INTERMEDIATE INFORMATION and LATE INFORMATION are represented by 0122110 and 0011122, respectively. We chose the width reductions so that the total reduction is constant across the 3 cases (in this example equal to 7); the starting conditional distribution of final demand is identical for the 3 cases, and final period width is same in all 3 cases. Presumably the firm has to invest in better forecasting systems to obtain the refined updates earlier in the horizon. We do not consider such costs in our analysis but only focus on the benefits.

Capacity Restrictions. We vary the capacity restrictions by changing the initial lower bound a_n for the

demand forecast. Higher values of a_n reflect tighter capacity restrictions. In the initial study, we consider $a_n = 0, a_n = 2$, and $a_n = 4$. We also considered $a_n = 3, a_n = 6$, and $a_n = 9$ while performing the tests on longer time horizons.

Mean Heuristic. Mean Heuristic (MH), which utilizes the mean forecast information, is intended to model the updated point forecasts of average demand that are often used in industry. In this case, in each period the manufacturer produces a unit if the mean of the forecasted demand cannot be reached by production in future periods. Thus, production takes place with n periods to go if the on-hand inventory

$$x < a_n + \lceil w_n/2 \rceil - n.$$

In all our experiments we assume that the distribution of future forecast bands is given by a uniform distribution. This implies that if the current forecast is given by $a_n = 2, w_n = 4$, and $w_{n-1} = 2$, then it is equally likely that the next forecast band would be $a_{n-1} = 2, w_{n-1} = 2$ or $a_{n-1} = 3, w_{n-1} = 2$ or $a_{n-1} = 4, w_{n-1} = 2$.

4.1. Heuristics Performance

In this section, we describe how the heuristics HLB, HUB, HCU, HCL, and MH perform compared to the optimal solution. The average error for HLB, HUB, HCU, HCL, and MH were 8.89%, 0.23%, 3.18%, 0.43%, and 12.13%, respectively, with standard deviations of 17.27%, 0.46%, 6.7%, 1.0%, and 15.81%, respectively, based on the 1,980 samples in our initial study. We computed the optimal expected costs using dynamic programming. Note that HUB and HCL perform much better in our set of experiments because both these heuristics tend to avoid overproduction. It is to be noted that overproduction in the first few periods could be quite expensive. Further, the performance of MH is poor compared to HUB, HLB, HCU, and HCL because it does not utilize any information about forecast evolution while planning production. In addition, it uses a point forecast (which is computed as a mean of the distribution in our implementation) and tries to produce to meet that demand in n periods. Our results show that cost differences from the optimal could be as large as 12% on average and even more under certain conditions.

Table 1 Average Error in Cost for HLB, HUB, HCU, HCL, and MH Under Different Capacity Conditions

a_n	Avg. HLB	Avg. HUB	Avg. HCU	Avg. HCL	Avg. MH
$a = 0$	14.08%	0.30%	6.72%	0.8%	18.23%
$a = 2$	8.50%	0.26%	2.29%	0.4%	10.97%
$a = 4$	4.10%	0.12%	0.52%	0.1%	7.20%

Table 2 Average Error in Cost for HLB, HUB, HCU, HCL, and MH Under Different Holding Costs

Holding cost	Avg. HLB	Avg. HUB	Avg. HCU	Avg. HCL	Avg. MH
$h = 0$	0.0%	0.0%	0.31%	1.2%	19.69%
$h = 2$	2.00%	0.28%	2.45%	0.54%	12.38%
$h = 4$	7.30%	0.31%	2.98%	0.23%	8.45%
$h = 8$	13.34%	0.32%	5.09%	0.11%	8.55%
$h = 12$	21.84%	0.23%	5.09%	0.03%	11.60%

Capacity Parameter. (1) We observe that the performance of both HUB and HLB improves with increase in a_n (see Table 1); this effect is more prominent for HLB. A higher value of a_n increases production requirements, and as a result HLB (which tends to produce more) performs better. Also when a_n is higher, there is a smaller chance that the algorithm will execute Step 1b.3, and this also contributes to better performance for both HUB and HLB. (2) The performance of both HCL and HCU improves under tighter capacity. HCU should be better because when capacity is tighter it is optimal to produce in more periods. HCL also performs better under tighter capacity because there is a smaller chance that the algorithm will execute Step 3. HCL. (3) The performance of MH improves with increasing a_n , which suggests that this heuristic is more effective when capacity is tight.

Cost Parameters. (1) The performance of HUB and HLB changes in different ways with increases in penalty and holding costs (see Tables 2 and 3). As penalty costs increase, the performance of HUB improves, while that of HLB gets worse (see Table 3). Underproduction hurts more when penalty costs are higher, which hurts HUB performance, while the expected additional costs for carrying inventory are lower, which helps HLB. (2) Increases in holding

Table 3 Average Error in Cost for HLB, HUB, HCU, HCL, and MH Under Different Penalty Costs

Penalty cost	Avg. HLB	Avg. HUB	Avg. HCU	Avg. HCL	Avg. MH
$\pi = 75$	17.16%	0.02%	1.18%	0.23%	9.19%
$\pi = 150$	6.38%	0.25%	4.70%	0.43%	7.56%
$\pi = 250$	3.16%	0.39%	3.66%	0.63%	19.63%

Table 4 Average Error in Cost for HLB, HUB, HCU, HCL, and MH for Different Time Horizons

Time Horizon	Avg. HLB	Avg. HUB	Avg. HCU	Avg. HCL	Avg. MH
$n = 4$	4.45%	0.02%	0.32%	0.04%	7.44%
$n = 8$	6.09%	0.24%	2.50%	0.43%	2.64%
$n = 12$	18.04%	0.37%	6.01%	0.73%	15.14%

costs reduce the effectiveness of HLB as expected but do not affect the performance of HUB in a significant manner (see Table 2). Note that when holding costs are zero both the heuristics are optimal. (3) The performance of HCU becomes worse while that of HCL improves with increases in holding cost, as one would expect. Similarly, the performance of HCU improves and that of HCL worsens with increasing penalty cost. (4) We find that the performance of MH first improves and then deteriorates with increases in holding and penalty costs. One explanation for this effect is that when penalty (holding) cost is very high (low), the optimal production quantity will be significantly larger than the mean. Similarly, when penalty (holding) cost is very low (high), then the optimal production may be lower than and away from the mean.

Time Horizon. We find that the performance of the three heuristics deteriorates with increase in time horizon. This is probably because there are likely to be more errors in the production decisions under longer time horizons. The cost increases in HLB, HCU, and MH are significant, whereas those in HCL and HUB are not as large (see Table 4). In all, the heuristics that produce less (HCL and HUB) are very close to optimal (less than 0.5%, on average) and perform much better than the heuristics that produce more (HCU

Table 5 Average Cost Difference Between EARLY, INTERMEDIATE, and LATE Information Updates Under Different Capacity Conditions

a_n	Avg. ($I - E$)	Avg. ($L - I$)	Avg. ($L - E$)
$a = 3$	4.39	13.72	18.11
$a = 6$	4.15	8.04	12.19
$a = 9$	1.97	3.78	5.76

and HLB). Further, the performance of HCL and HUB is orders of magnitude better than MH, which utilizes only point forecasts.

4.2. Timing of Information Updates

To test the effect of the timing of additional information availability during forecast updates, we conducted another study with 1,080 problems. The time horizon of 8, 12, 16, 18, and 20 were considered; holding cost was set to 0, 2, 8, and 12; salvage cost was chosen to be 0, 25, and 45; penalty costs of 75 and 150 were considered; and the capacity available was varied with a_n equal to 3, 6, and 9. For each of these parameter settings 3 different information update scenarios were considered: EARLY, INTERMEDIATE, and LATE (described at the beginning of §4). Their optimal costs are represented by E , I , and L , respectively.

The average values of the differences $I - E$, $L - I$, and $L - E$ were 3.50, 8.51, and 12.02, respectively. We find that the availability of more refined forecasts early on reduces the optimal total cost. This is intuitive because having the updates early reduces the demand uncertainty. Although $I - E$ as well as $L - I$ are not very large (on average), we find that $L - I$ is greater than $I - E$. From a managerial standpoint, this indicates that if it is not economically viable to obtain early information updates, the firm should not lose hope but strive to achieve that information during the horizon.

Capacity Parameter. As the capacity becomes tighter, the benefits due to early information reduce (see Table 5). This result is intuitive in that under tighter capacity it is likely to be optimal to produce in a greater number of periods irrespective of information availability.

Cost Parameters. (1) As the holding cost increases

Table 6 Average Cost Difference Between EARLY, INTERMEDIATE, and LATE Information Updates Under Different Holding Costs

Holding cost	Avg. ($I - E$)	Avg. ($L - I$)	Avg. ($L - E$)
$h = 0$	0.11	12.59	12.70
$h = 2$	8.89	12.78	21.67
$h = 8$	4.03	6.10	10.13
$h = 12$	0.99	2.58	3.58

Table 7 Average Cost Difference Between EARLY, INTERMEDIATE, and LATE Information Updates Under Different Penalty Costs

Penalty Cost	Avg. ($I - E$)	Avg. ($L - I$)	Avg. ($L - E$)
$\pi = 75$	0.59	3.43	4.02
$\pi = 150$	6.41	13.60	20.02

we observe that the differences $I - E$ and $L - I$ first increase and then decrease (see Table 6). This implies that earlier updates are more beneficial when holding costs are neither too high nor too low. Because the holding cost plays a crucial role in the production decision of each period (particularly in early periods), it is expected that very low or very high values bias the production decision more than the demand uncertainty. As a result, in those cases benefits of more refined forecasts are lower. A managerial implication is that for products with variable demand and high holding costs (such as high tech) the benefits of obtaining early updates may not be substantial because even if the information is available early, it may be optimal to produce only later in the horizon. (2) We also find that when holding costs are lower, $L - I$ is significantly higher than $I - E$. An explanation for this effect is that at lower holding costs the wrong production decisions, due to the absence of information earlier on, have a limited effect because future periods are available for adjustment after the major updates are received. However, $L - I$ remains high because there may not be enough time remaining to adjust production. (3) As expected, we find that with increases in penalty costs, the benefit of obtaining early refinements in forecasts increases (see Table 7).

Time Horizon. (1) We find that the benefits of obtaining earlier updates are greater when the time horizons are longer (see Table 8). In our experiments, a

Table 8 Average Cost Difference Between EARLY, INTERMEDIATE, and LATE Information Updates for Different Time Horizons

Time Horizon	Avg. $(I - E)$	Avg. $(L - I)$	Avg. $(L - E)$
$n = 8$	1.70	1.90	3.61
$n = 12$	3.31	5.85	9.17
$n = 16$	4.01	11.28	15.29
$n = 18$	4.03	11.77	15.81
$n = 20$	4.45	11.77	16.23

longer horizon means a greater level of demand uncertainty in the beginning periods. As a result, the value of additional information is higher. (2) We also notice that the marginal benefit due to earlier updates decreases with the number of periods in the horizon. This can be explained by the fact that longer horizons allow more time for adjustment when there are later updates; as a result, the marginal benefit of early information is not as substantial.

5. Multiple Unit Capacity: No Holding-Cost Case

In this section we consider the multiple capacity case where the firm has the flexibility to produce up to C units in each period of time, and focus on the no holding-cost case. We extend our algorithm HUB/HLB for the no holding-cost case (recall that they are equivalent in this case) and prove that the algorithm is optimal in this case. These algorithms can be extended in an equivalent way for the holding-cost case.

First, all the possible values of a_1 and their probabilities are generated using the F_t for all t from n to 1 and the values of a_n and w_n . Computationally, the probability values are easy to calculate because all values are discrete. The heuristic continues as follows:

Step 1. For $a_1 \in \{a_n, a_n + 1, \dots, (a_n + w_n - w_1)\}$,

Step 1a. Calculate x^* taking current inventory of x_n into account. Specifically,

$$x^* = \underset{a_1 \leq y \leq a_1 + w_1}{\operatorname{argmin}} \{ \tilde{V}_0(y, a_1, w_1) + c(y - x_n)^+ \},$$

where $\tilde{V}_0(x, a_1, w_1)$ is the single period cost to go if inventory is x , and demand can take on the discrete values between a_1 and $a_1 + w_1$ with a probability distribution given by $F_1(a_1)$.

Step 1b. For $l = 1, 2, \dots, C$, the following 3 cases are considered:

Step 1b.1. If $x_n + l \geq x^*$, then $\lambda_{a_1}^l = c + \tilde{V}_0(x_n + l, a_1, w_1) - \tilde{V}_0(x_n + l - 1, a_1, w_1)$.

Step 1b.2. If $x_n + l + C(n - 1) \leq x^*$, then $\lambda_{a_1}^l = c + \tilde{V}_0(x_n + l + C(n - 1), a_1, w_1) - \tilde{V}_0(x_n + l - 1 + C(n - 1), a_1, w_1)$.

Step 1b.3. If $x_n + l < x^* < x_n + C(n - 1) + l$, $\lambda_{a_1}^l = 0$.

Step 2. Let $p_{a_1|a_n}$ be the conditional probabilities of each a_1 given the current a_n . Let $\lambda_n^l = \sum_{a_1=a_n}^{a_n+w_n-w_1} p_{a_1|a_n} \lambda_{a_1}^l$ be the weighted average value of $\lambda_{a_1}^l$.

Step 3. If there are one or more negative λ_n^l s, produce q such that $\lambda_n^q = \max \lambda_n^l: \lambda_n^l < 0$; if all $\lambda_n^l \geq 0$, don't produce.

THEOREM 5.1. *This heuristic is optimal when the holding cost equals zero.*

6. Conclusions

We present a model that captures forecast evolution information while making production decisions in a capacitated environment. We consider a forecast evolution model that is defined by a band (with upper and lower bounds) that captures the uncertainty in the forecast. As time moves forward, we assume the next forecast has a smaller width (representing a better forecast), and that the new band can lie anywhere inside the old band. This models a forecasting process that gets refined over time as new information arrives. We consider a manufacturing firm utilizing this forecast that has a fixed capacity in each period and needs to decide in which periods it should produce, taking into account expected production, holding, salvage, and stock-out costs. Demand occurs in the final period, modeling a seasonal product with terminal demand. We prove the existence of inventory threshold levels in each period below which the firm should produce and above which it should not. We provide several simple heuristics for the holding-cost case, demonstrate that two of them are optimal in the zero-holding-cost case, and characterize their relationship to the optimal production policy. Our computational study shows that utilizing the additional information in the bands provides better solutions

than a heuristic that utilizes a point forecast. Further, we find that two of the heuristics are very close to optimal on average (less than 0.5% away).

We empirically observe that early information updates in the forecasting process lead to decreased costs. On average, we find that majority of the benefits can be realized if the information updates are received by the middle of the planning horizon. Little additional benefit is obtained by improving forecasts early in the process. Further, we find that information updates are most useful when holding costs are neither too high nor too low and when the capacity is not too tight. Finally, we find that although the benefits of early information are greater in periods with longer horizons, the increases in benefits diminish as horizons get very long.

This paper presents a model for capturing forecast updates and integrating it with production planning for a product with terminal demand. There are certain limitations in this model. First, we assume that forecasts get better or at least are as good in future periods. Although this assumption is reasonable under most conditions, there may be situations where this condition may be violated. Second, we present a model where only one product is being produced in a capacitated setting. A multiproduct system would be more appropriate in a general setting; however, it will be more complex to handle. Finally, the general case where demand can occur in any period is more complicated and is a topic for future research.

Acknowledgments

The authors would like to thank the editor and referees for their valuable feedback and comments on earlier drafts of this paper. In addition, seminar participants at the INFORMS conferences in Philadelphia and Salt Lake City, as well as the M&SOM conference at Ann Arbor, provided useful feedback. The research of the first author was supported in part by NSF Contract DMI-9732795, and the research of the second author was supported in part by NSF CAREER award DMI-9984252.

Appendix

PROOF OF LEMMA 3.1. Let $V1(\xi) = s\{(x + 2 - \xi)^+ - (x + 1 - \xi)^+\} + \pi\{(\xi - x - 2)^+ - (\xi - x - 1)^+\}$ and $V2(\xi) = s\{(x + 1 - \xi)^+ - (x - \xi)^+\} + \pi\{(\xi - x - 1)^+ - (\xi - x)^+\}$. Clearly, if $\xi \geq x + 2$, then $V1(\xi) = V2(\xi) = -\pi$; if $\xi \leq x$, then $V1(\xi) = V2(\xi) = s$; and if $\xi = x + 1$, then $V1(\xi) = s$, and $V2(\xi) = -\pi$. Thus, in all cases $V1(\xi) \geq V2(\xi)$. This implies that $\sum_{\xi} p_{\xi} V1(\xi) \geq \sum_{\xi} p_{\xi} V2(\xi)$ for any set of values for ξ . Therefore, $\sum_{a_0=a_1}^{a_1+w_1} p_{a_0} V1(a_0) \geq \sum_{a_0=a_1}^{a_1+w_1} p_{a_0} V2(a_0)$. So, we conclude

that $\mathbf{E}_{a_0}\{V_0(x + 2, a_0, 0) - V_0(x + 1, a_0, 0)\} \geq \mathbf{E}_{a_0}\{V_0(x + 1, a_0, 0) - V_0(x, a_0, 0)\}$. \square

PROOF OF THEOREM 3.2. The optimal recursion function is given by

$$V_{t+1}(x_{t+1}, a_{t+1}, w_{t+1}) \\ = \min \begin{cases} c + h(x_{t+1} + 1) + \mathbf{E}_{a_t} V_t(x_{t+1} + 1, a_t, w_t) \\ hx_{t+1} + \mathbf{E}_{a_t} V_t(x_{t+1}, a_t, w_t). \end{cases}$$

For the remainder of this proof, we suppress the subscripts on the expectations. To prove the theorem, it is sufficient to show that $c + h + \mathbf{E}V_t(x + 1, a_t, w_t) - \mathbf{E}V_t(x, a_t, w_t)$ is monotone increasing in x . The proof proceeds through induction on the number of periods to go.

From Lemma 3.1, we know the result holds for $t = 0$. Now, assume that the result holds for $1, 2, \dots, t - 1$, and consider the following cases at time t .

Case 1. $V_t(x + 1, a_t, w_t) = c + h(x + 2) + \mathbf{E}V_{t-1}(x + 2, a_{t-1}, w_{t-1})$ and $V_t(x, a_t, w_t) = c + h(x + 1) + \mathbf{E}V_{t-1}(x + 1, a_{t-1}, w_{t-1})$. This means that

$$V_t(x + 1, a_t, w_t) - V_t(x, a_t, w_t) \\ = h + \mathbf{E}\{V_{t-1}(x + 2, a_{t-1}, w_{t-1}) - V_{t-1}(x + 1, a_{t-1}, w_{t-1})\},$$

so by the induction hypothesis, monotonicity of the differences holds.

Case 2. $V_t(x + 1, a_t, w_t) = h(x + 1) + \mathbf{E}V_{t-1}(x + 1, a_{t-1}, w_{t-1})$ and $V_t(x, a_t, w_t) = c + h(x + 1) + \mathbf{E}V_{t-1}(x + 1, a_{t-1}, w_{t-1})$. This means that $V_t(x + 1, a_t, w_t) - V_t(x, a_t, w_t) = -c$, so the difference is constant and, thus, monotonicity trivially holds.

Case 3. $V_t(x + 1, a_t, w_t) = h(x + 1) + \mathbf{E}V_{t-1}(x + 1, a_{t-1}, w_{t-1})$ and $V_t(x, a_t, w_t) = h(x) + \mathbf{E}V_{t-1}(x, a_{t-1}, w_{t-1})$. This means that

$$V_t(x + 1, a_t, w_t) - V_t(x, a_t, w_t) \\ = h + \mathbf{E}\{V_{t-1}(x + 1, a_{t-1}, w_{t-1}) - V_{t-1}(x, a_{t-1}, w_{t-1})\},$$

so by the induction argument, monotonicity of the differences holds.

Case 4. $V_t(x + 1, a_t, w_t) = c + h(x + 2) + \mathbf{E}V_{t-1}(x + 2, a_{t-1}, w_{t-1})$ and $V_t(x, a_t, w_t) = h(x) + \mathbf{E}V_{t-1}(x, a_{t-1}, w_{t-1})$. In this case,

$$V_t(x + 1, a_t, w_t) - V_t(x, a_t, w_t) \\ = c + 2h + \mathbf{E}\{V_{t-1}(x + 2, a_{t-1}, w_{t-1}) - V_{t-1}(x, a_{t-1}, w_{t-1})\} \\ = c + 2h + \mathbf{E}\{V_{t-1}(x + 2, a_{t-1}, w_{t-1}) - V_{t-1}(x + 1, a_{t-1}, w_{t-1}) \\ + V_{t-1}(x + 1, a_{t-1}, w_{t-1}) - V_{t-1}(x, a_{t-1}, w_{t-1})\}.$$

Once again, monotonicity follows from the induction hypothesis. \square

PROOF OF PROPERTY 3.4. To see this, consider the three conditions for each value of a_i in the algorithm.

1. For the first case where $x_n \geq x^*$, the bound is met at equality for (i). For (ii) we use the convexity of the value function, noting that $\tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1) \leq \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1)$.
2. For the second case where $x_n + n \leq x^*$, we use the convexity

of the value function to show $\tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1) \geq \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)$ for (i). For (ii) the bound is met at equality.

3. For the final case, we note that $\tilde{V}_0(x_n + 1, a_1, w_1) + c < \tilde{V}_0(x_n, a_1, w_1)$ because $x^* = \operatorname{argmin}_{a_1 \leq y \leq a_1 + w_1} \tilde{V}_0(y, a_1, w_1) + c(y - x_n)$ and $\lambda_{a_1} = 0$, so (i) must be true. Similarly, because $x_n + n \geq x^*$ and value function is convex, (ii) must be true. \square

PROOF OF PROPERTY 3.5. Note that this is a special case of property (iii) in the proof of theorem 5.1 where $C = 1$ and $l = 1$. So, proof follows directly from that. A detailed proof is available from the authors. \square

PROOF OF PROPERTY 3.8. Consider the three conditions for each value of a_1 in the heuristic.

1. For the first case $x_n \geq x^*$, the result holds at equality for (i). For (ii) from Lemma 3.1, we know that $\tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1) \leq \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1)$.

2. For the second case $x_n + n \leq x^*$, for (i) $\tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1) \geq \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)$ based on Lemma 3.1. Thus, $c + nh + \tilde{V}_0(x_n + n, a_1, w_1) - \tilde{V}_0(x_n + n - 1, a_1, w_1) \geq c + nh + \tilde{V}_0(x_n + 1, a_1, w_1) - \tilde{V}_0(x_n, a_1, w_1)$. For (ii) this holds at equality.

3. For the final case, for (i) we note that $\tilde{V}_0(x_n + 1, a_1, w_1) + c + h \leq \tilde{V}_0(x_n, a_1, w_1)$ because of the definition of x^* . Because 1b.3.ii simplifies to $(n - (x^* - x_n))(x^* - x_n)h$, which is greater than $(n - 1)h$ since $0 < (x^* - x_n) < n$, we know that $\tilde{\lambda}_n \geq (n - 1)h$ in this case, so (i) must be true. For (ii), we note that $\tilde{V}_0(x_n + n - 1, a_1, w_1) \leq c + nh + \tilde{V}_0(x_n + n, a_1, w_1)$. Because 1b.3.i simplifies to $(n - 1)h - (n - (x^* - x_n))(x^* - x_n)h$ and $0 < (x^* - x_n) < n$, we have that $\underline{\lambda}_n \leq 0$, so (ii) must be true. \square

PROOF OF PROPERTY 3.9. Consider (i) of the property. By definition we have,

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + (k-1)h + V_{n-k+1}(x+1, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \\ & = E_{a_{n-k+1}|a_n}\{c + (k-1)h + \min(A, B) - \min(C, D)\}, \end{aligned}$$

where A, B, C , and D are defined as follows:

$$A = E_{a_{n-k}|a_{n-k+1}}\{c + (x+2)h + V_{n-k}(x+2, a_{n-k}, w_{n-k})\},$$

$$B = E_{a_{n-k}|a_{n-k+1}}\{(x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\},$$

$$C = E_{a_{n-k}|a_{n-k+1}}\{c + (x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\},$$

$$D = E_{a_{n-k}|a_{n-k+1}}\{xh + V_{n-k}(x, a_{n-k}, w_{n-k})\}.$$

We consider the following cases.

Case 1. $A \leq B, C \leq D$. This means that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + (k-1)h + V_{n-k+1}(x+1, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \\ & = E_{a_{n-k+1}|a_n}\{c + (k-1)h \\ & \quad + E_{a_{n-k}|a_{n-k+1}}\{(x+2)h + V_{n-k}(x+2, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{(x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\}\} \end{aligned}$$

$$\begin{aligned} & = E_{a_{n-k}|a_n}\{c + kh + V_{n-k}(x+2, a_{n-k}, w_{n-k}) \\ & \quad - V_{n-k}(x+1, a_{n-k}, w_{n-k})\} \\ & \geq E_{a_{n-k}|a_n}\{c + kh + V_{n-k}(x+1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})\}, \end{aligned}$$

where the last step is due to the monotonicity of $V_{n-k}(x+1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})$ with respect to x .

Case 2. $A \leq B, D \leq C$. This means that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + (k-1)h + V_{n-k+1}(x+1, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \\ & = E_{a_{n-k+1}|a_n}\{c + (k-1)h \\ & \quad + E_{a_{n-k}|a_{n-k+1}}\{c + (x+2)h + V_{n-k}(x+2, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{xh + V_{n-k}(x, a_{n-k}, w_{n-k})\}\} \\ & \geq E_{a_{n-k+1}|a_n}\{c + (k-1)h \\ & \quad + E_{a_{n-k}|a_{n-k+1}}\{c + (x+2)h + V_{n-k}(x+2, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{c + (x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\}\} \\ & \geq c + kh + E_{a_{n-k}|a_n}\{V_{n-k}(x+2, a_{n-k}, w_{n-k}) \\ & \quad - V_{n-k}(x+1, a_{n-k}, w_{n-k})\} \\ & \geq E_{a_{n-k}|a_n}\{c + kh + V_{n-k}(x+1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})\}. \end{aligned}$$

Case 3. $B \leq A, C \leq D$. This means that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + (k-1)h + V_{n-k+1}(x+1, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \\ & = c + (k-1)h \\ & \quad + E_{a_{n-k+1}|a_n}\{E_{a_{n-k}|a_{n-k+1}}\{(x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\} \\ & \quad - c - E_{a_{n-k}|a_{n-k+1}}\{(x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\}\} \\ & = (k-1)h. \end{aligned}$$

Case 4. $B \leq A, D \leq C$. This means that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + (k-1)h + V_{n-k+1}(x+1, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x, a_{n-k+1}, w_{n-k+1})\} \\ & = c + (k-1)h \\ & \quad + E_{a_{n-k+1}|a_n}\{E_{a_{n-k}|a_{n-k+1}}\{(x+1)h + V_{n-k}(x+1, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{xh + V_{n-k}(x, a_{n-k}, w_{n-k})\}\} \\ & = c + kh + E_{a_{n-k}|a_n}\{V_{n-k}(x+1, a_{n-k}, w_{n-k}) - V_{n-k}(x, a_{n-k}, w_{n-k})\}. \end{aligned}$$

This proves (i). The second part of the property (ii) can be proved with a similar approach, hence, the proof has been omitted. \square

PROOF OF THEOREM 5.1. To prove this result, we need several properties (i)–(iv).

(i) $\lambda_n^l \geq E_{a_1|a_n}\{c + \tilde{V}_0(x_n + l, a_1, w_1) - \tilde{V}_0(x_n + l - 1, a_1, w_1)\}$; (ii) λ_n^l

$$\leq E_{a_1|a_n}\{c + \tilde{V}_0(x_n + l + C(n-1), a_1, w_1) - \tilde{V}_0(x_n + l - 1 + C(n-1), a_1, w_1)\}.$$

Proof:

(i) To see this, consider the three conditions for each value of l and a_1 in the algorithm.

1. For the first case, the bound is met at equality.
2. For the second case, we use the monotonicity of the differences in the value function.

3. For the final case, we note that $\tilde{V}_0(x_n + l, a_1, w_1) + c < \tilde{V}_0(x_n + l - 1, a_1, w_1)$ because $x^* = \operatorname{argmin}_{a_1 \leq y \leq a_1 + w_1} \{\tilde{V}_0(y, a_1, w_1) + c(y - x_n)\}$. Because $\lambda_n^l = 0$, the result follows.

(ii) To see this, consider the three conditions considered for each value of l and a_1 in the algorithm.

1. For the first case, we use the convexity of the value function.
2. For the second case, the bound is met at equality.
3. For the final case, we note that $\tilde{V}_0(x_n + l + C(n-1), a_1, w_1) + c > \tilde{V}_0(x_n + l - 1 + C(n-1), a_1, w_1)$ because $x^* = \operatorname{argmin}_{a_1 \leq y \leq a_1 + w_1} \{\tilde{V}_0(y) + c(y - x_n)\}$. Because $\lambda_n^l = 0$, the result follows.

Next we prove the following properties (iii) and (iv).

(iii) If $E_{a_{n-k}|a_n}\{c + V_{n-k}(x + l + C(k-1), a_{n-k}, w_{n-k}) - V_{n-k}(x + l - 1 + C(k-1), a_{n-k}, w_{n-k})\} \geq 0$, then $E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + l + C(k-2), a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x + l - 1 + C(k-2), a_{n-k+1}, w_{n-k+1})\} \geq 0$; (iv) if $E_{a_{n-k}|a_n}\{c + V_{n-k}(x + l, a_{n-k}, w_{n-k}) - V_{n-k}(x + l - 1, a_{n-k}, w_{n-k})\} \leq 0$, then $E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + l, a_{n-k+1}, w_{n-k+1}) - V_{n-k+1}(x + l - 1, a_{n-k+1}, w_{n-k+1})\} \leq 0$.

Proof:

(iii)

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + l + C(k-2), a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x + l - 1 + C(k-2), a_{n-k+1}, w_{n-k+1})\} \\ & = E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\}, \end{aligned}$$

where A_i and B_j are defined as follows:

$$\begin{aligned} A_i &= E_{a_{n-k}|a_{n-k+1}}\{ci + V_{n-k}(x + l + i + C(k-2), a_{n-k}, w_{n-k})\} \\ B_j &= E_{a_{n-k}|a_{n-k+1}}\{cj + V_{n-k}(x + l - 1 + j + C(k-2), a_{n-k}, w_{n-k})\}. \end{aligned}$$

Let i^* be the minimizing i value, and consider the following cases:

• $i^* < C$. In this case, we can set $j = i + 1$ and conclude that:

$$\min_{0 \leq j \leq C} B_j \leq E_{a_{n-k}|a_{n-k+1}}\{c(i+1) + V_{n-k}(x + l + i + C(k-2), a_{n-k}, w_{n-k})\}$$

and, hence,

$$\min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j \geq -c,$$

so

$$E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\} \geq 0.$$

• $i^* = C$. In this case, we can substitute C for both i and j , and we conclude that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\} \\ & = E_{a_{n-k+1}|a_n}\left\{c + A_C - \min_{0 \leq j \leq C} B_j\right\} \\ & \geq E_{a_{n-k+1}|a_n}\{c + A_C - B_C\} \\ & = c + E_{a_{n-k+1}|a_n}\{E_{a_{n-k}|a_{n-k+1}}\{cC + V_{n-k}(x + l + (k-1)C, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{cC + V_{n-k}(x + l - 1 \\ & \quad \quad \quad + (k-1)C, a_{n-k}, w_{n-k})\}\} \\ & = E_{a_{n-k}|a_n}\{c + V_{n-k}(x + l + C(k-1), a_{n-k}, w_{n-k}) \\ & \quad - V_{n-k}(x + l - 1 + C(k-1), a_{n-k}, w_{n-k})\}. \end{aligned}$$

(iv)

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\{c + V_{n-k+1}(x + l, a_{n-k+1}, w_{n-k+1}) \\ & \quad - V_{n-k+1}(x + l - 1, a_{n-k+1}, w_{n-k+1})\} \\ & = E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\}, \end{aligned}$$

where A_i and B_j are defined as follows:

$$\begin{aligned} A_i &= E_{a_{n-k}|a_{n-k+1}}\{ci + V_{n-k}(x + l + i, a_{n-k}, w_{n-k})\} \\ B_j &= E_{a_{n-k}|a_{n-k+1}}\{cj + V_{n-k}(x + l - 1 + j, a_{n-k}, w_{n-k})\}. \end{aligned}$$

Let j^* be the minimizing j value, and consider the following cases:

• $j^* > 0$. In this case, we can set $i = j - 1$ and conclude that

$$\min_{0 \leq i \leq C} A_i \leq E_{a_{n-k}|a_{n-k+1}}\{c(j-1) + V_{n-k}(x + l - 1 + j, a_{n-k}, w_{n-k})\}$$

and, hence,

$$\min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j \leq -c,$$

so

$$E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\} \leq 0.$$

• $j^* = 0$. In this case, we can substitute 0 for both i and j , and we conclude that

$$\begin{aligned} & E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - \min_{0 \leq j \leq C} B_j\right\} \\ & = E_{a_{n-k+1}|a_n}\left\{c + \min_{0 \leq i \leq C} A_i - B_0\right\} \\ & \leq E_{a_{n-k+1}|a_n}\{c + A_0 - B_0\} \\ & = c + E_{a_{n-k+1}|a_n}\{E_{a_{n-k}|a_{n-k+1}}\{V_{n-k}(x + l, a_{n-k}, w_{n-k})\} \\ & \quad - E_{a_{n-k}|a_{n-k+1}}\{V_{n-k}(x + l - 1, a_{n-k}, w_{n-k})\}\} \end{aligned}$$

$$= E_{a_{n-k}|a_n}\{c + V_{n-k}(x + l, a_{n-k}, w_{n-k}) - V_{n-k}(x + l - 1, a_{n-k}, w_{n-k})\}.$$

Repeated applications of (iii) and (iv) result in (v) and (vi), respectively: (v). If $E_{a_1|a_n}\{c + \tilde{V}_0(x + l + C(n - 1), a_1, w_1) - \tilde{V}_0(x + l - 1 + C(n - 1), a_1, w_1)\} \geq 0$, then $E_{a_{n-1}|a_n}\{c + V_{n-1}(x + l, a_{n-1}, w_{n-1}) - V_{n-1}(x + l - 1, a_{n-1}, w_{n-1})\} \geq 0$; (vi) if $E_{a_1|a_n}\{c + \tilde{V}_0(x + l, a_1, w_1) - \tilde{V}_0(x + l - 1, a_1, w_1)\} \leq 0$, then $E_{a_{n-1}|a_n}\{c + V_{n-1}(x + l, a_{n-1}, w_{n-1}) - V_{n-1}(x + l - 1, a_{n-1}, w_{n-1})\} \leq 0$.

Combining (v) and (vi) with (i) and (ii) proves that the heuristic is optimal. \square

References

- Agarwal, N., S. Smith, A. A. Tsay. 1999. Multivendor sourcing in a retail supply chain. Working paper, Santa Clara University, CA.
- Aviv, Y., A. Federgruen. 1997. Stochastic inventory models with limited production capacity and periodically varying parameters. *Probab. Engrg. Informational Sci.* **11** 107–135.
- Azoury, K. S. 1985. Bayes solution to dynamic inventory models under unknown demand distribution. *Management Sci.* **31**(9) 1150–1160.
- Bitran, G. R., E. A. Haas, H. Matsuo. 1986. Production planning of style goods with high setup costs and forecast revisions. *Oper. Res.* **34**(2) 226–236.
- Buckley, S., R. Jayaraman, G. Lin, N. Nayak, R. Srinivasan, J. M. Swaminathan, M. Taubenblatt. 1995. Responding to demand uncertainty through supplier flexibility and safety stocks. *INFORMS*, Los Angeles.
- Federgruen, A., P. Zipkin. 1986. An inventory model with limited production capacity and uncertain demands. I: The average cost criterion. *Math. Oper. Res.* **11**(2) 193–207.
- Fisher, M. L., A. Raman. 1996. Reducing the cost of demand uncertainty through accurate response to early sales. *Oper. Res.* **44**(1) 87–99.
- Gullu, R. 1996. On the value of information in dynamic production/inventory problems under forecast evolution. *Naval Res. Logist.* **43** 289–303.
- Gavirneni, S., R. Kapuscinski, S. Tayur. 1999. Value of information in capacitated supply chains. *Management Sci.* **45**(1) 16–24.
- Gurnani, H., C. S. Tang. 1999. Optimal ordering decisions with uncertain cost and demand forecast updating. *Management Sci.* **45**(10) 1456–1462.
- Hausman, W. H. 1969. Sequential decision problems: A model to exploit existing forecasters. *Management Sci.* **16**(2) B93–B110.
- , R. Peterson. 1972. Multiproduct production scheduling for style goods with limited capacity, forecast revisions, and terminal delivery. *Management Sci.* **18**(7) 370–383.
- Heath, D. C., P. L. Jackson. 1994. Modeling the evolution of demand forecasts with application to safety stock analysis in production/distribution systems. *IIE Trans.* **26**(3) 17–30.
- Kapuscinski, R., S. Tayur. 1998. A capacitated production—inventory model with periodic demand. *Oper. Res.* **46**(6) 899–911.
- Lee, H. L., J. Whang. 1998. Value of postponement. Tech Ho, Chris Tang, eds. *Product Variety Management: Research Advances*. Kluwer Publishers, Norwell, MA, 65–84.
- , V. Padmanabhan, J. Whang. 1997. Information distortion in a supply chain: The bull whip effect. *Management Sci.* **43**(4) 546–548.
- Morton, T. E., D. W. Pentico. 1995. The finite horizon nonstationary stochastic inventory problem: Near-myopic bounds, heuristics, testing. *Management Sci.* **41**(2) 334–343.
- Nurani, R. K., R. Akella, J. G. Shanthikumar. 1994. Effects of using future forecast information on myopic policies. *Proc. 1994 M&SOM*. Pittsburgh, PA.
- Scarf, H. 1959. Bayes solutions of the statistical inventory problem. *Ann. Math. Statist.* **30**(2) 490–508.
- Simchi-Levi, D., P. Kaminsky, E. Simchi-Levi. 2000. *Designing and Managing the Supply Chain*. Irwin-McGraw Hill, NY.
- Swaminathan, J. M., J. G. Shanthikumar. 1999. Supplier diversification: The effect of discrete demand. *Oper. Res. Lett.* **24**(5) 213–221.
- Toktay, L. B., L. M. Wein. 1999. Analysis of a forecasting production inventory system with stationary demand. Working paper, INSEAD, France.

The consulting Senior Editor for this manuscript was Hau Lee. This manuscript was received on August 20, 1999, and was with the authors 151 days for 3 revisions. The average review cycle time was 86 days.