All-or-Nothing Ordering under a Capacity Constraint and Forecasts of Stationary Demand

Guillermo Gallego
IEOR Department, Columbia University
500 West 120th Street, New York, NY 10027, USA

and

L. Beril Toktay
INSEAD
77305 Fontainebleau Cedex, France

Abstract

We consider a supplier who faces stationary demand and uses dynamically updated forecast information to place orders to an upstream distributor. The order size is restricted and each order entails a high fixed cost. We characterize the form of the optimal policy in the special case where the fixed cost of ordering is high enough to warrant all-or-nothing ordering in each period.

November 2003
1 Introduction

We consider a single-item, periodic-review inventory control problem in the presence of (correlated) stationary demand, a capacity constraint, fixed ordering costs and a delivery lag. Demand forecasts are available for a fixed number of periods into the future and are updated in each period. The goal is to minimize the total expected ordering, inventory holding and backorder costs over a finite time horizon.

This research is related to two streams of literature. The first incorporates forecast information in inventory control decisions, typically considering the joint forecasting and inventory control problem in an uncapacitated environment (Veinott 1965, Johnson and Thompson 1975, Miller 1986, Lovejoy 1992). Graves et al. (1986, 1998) and Heath and Jackson (1994) envision decentralized forecasting and inventory management processes, where forecast updates are exogenous inputs to the production planning/inventory control function. They independently propose a general characterization of a forecast update process, dubbed the Martingale Model of Forecast Evolution (MMFE) by Heath and Jackson. Güllü (1996) uses the MMFE in analyzing the performance of a capacitated single-product system with uncorrelated demands for which one-period ahead forecasts are available. Toktay and Wein (2001) also use the MMFE to model the forecast update process in a single-item capacitated make-to-stock production-inventory system with no setup cost, and prove that the optimal production policy is a state-dependent modified order-up-to policy with respect to forecast-corrected inventory. A related paper is Buzacott and Shanthikumar (1994), who analyze a capacitated continuous-time MRP system with iid demands that are known exactly over a fixed time horizon. Karaesmen et al. (2002) develop a dynamic programming formulation of a discretized, Markovian version of the Buzacott-Shanthikumar model with unit demand and production in each period, and analyze the impact of capacity on the value of advance demand information. Gallego and Özer (2001) prove the optimality of state-dependent $(s, S)$ policies in the single-item uncapacitated inventory control problem with setup cost and iid demands that are progressively revealed. In this paper, we assume that the exogenous forecast up-
date process is modelled according to the Martingale Model of Forecast Evolution described in §2.1.

The second stream of research concerns the characterization of the optimal policy in a single-item, periodic-review inventory control setting in the presence of iid demand and fixed ordering costs. While the uncapacitated version of this problem was solved in the seminal paper of Scarf (1960), the capacitated version has not been fully characterized to date. The most complete characterization appears in Gallego and Scheller-Wolf (2000), where it is shown that the optimal capacitated policy has an \((s, S)\)-like structure.

Two recent papers consider a special case of the capacitated problem where all orders are constrained to be full-capacity orders (e.g., full truckloads); this assumption is justified in cases with very high fixed ordering cost. Gallego and Toktay (2003) show that the optimal policy is a threshold policy, with respect to the inventory position, for a class of cost-to-go functions that includes the class of convex functions. Özer and Wei (2003) extend this result to a setting where independent demands are progressively revealed, that is, forecast updates are always non-negative, and show that the optimal policy is a threshold policy with respect to the inventory position modified by the observed part of lead time demand. In addition, Özer and Wei (2003) do an extensive analysis of the impact of capacity level and fixed cost on the threshold level and the optimal cost under advance demand information. We do not attempt to duplicate such an analysis here, but prove that the structural result extends to the case of correlated demand and forecast updates that allow for both negative and positive corrections, as modelled by the MMFE.

## 2 The Model

We describe the Martingale Model of Forecast Evolution in §2.1. In §2.2, we formulate the inventory control problem of minimizing the total expected ordering, inventory holding and backorder costs over a finite horizon for a supplier receiving an exogeneous sequence of forecast updates about future demands.
2.1 The Martingale Model of Forecast Evolution

Let $D_n$ denote the demand in period $n$, where $\{D_n\}$ is stationary with $E[D_n] = \lambda$. Define $H$ to be the largest integer $i$ such that $\text{Cov}(D_n, D_{n+i}) \neq 0$. This is the forecast horizon over which non-trivial forecasts are available. Denote by $F_{n,n+i}$, $i = 0, 1, \ldots$ the forecast of period $n + i$ demand made in period $n$ after current demand is observed; this construct implies that $F_{n,n} = D_n$. For $i > H$, $F_{n,n+i} = \lambda$. Let $\varepsilon_{n,n+i} = F_{n,n+i} - F_{n-1,n+i}$ be the forecast update for period $n + i$ demand recorded in period $n$. Then $\varepsilon_n = (\varepsilon_{n,n}, \varepsilon_{n,n+1}, \ldots, \varepsilon_{n,n+H})$ is the vector of forecast updates over the forecast horizon $H$ recorded in period $n$. The Martingale Model of Forecast Evolution is a descriptive model that characterizes the sequence of forecast update vectors $\{\varepsilon_n\}$. It posits that if forecasts are unbiased and forecast updates are uncorrelated (see Heath and Jackson for conditions giving rise to this structure), then $\{\varepsilon_n\}_{n=-\infty}^{\infty}$ forms an independent and identically distributed sequence with $\varepsilon_n \sim N(0, \Sigma)$. Note that the MMFE is a descriptive model, and not a forecasting tool. To characterize an existing forecast update process, past forecast update data would be used to estimate the components $\sigma_{ij}, 0 \leq i, j \leq H$ of $\Sigma$. In this manner, the MMFE can be used to model forecasts arising from time-series forecasting methods, expert judgement, advance order information, etc. In the remainder of this paper, we will assume that the assumptions of the MMFE are satisfied.

The forecast $F_{n,n+i}$ can be reconstructed in terms of past forecast updates: $F_{n,n+i} = \lambda + \varepsilon_{n+i-H,n+i} + \ldots + \varepsilon_{n,n+i}$.

2.2 The Inventory Control Model

A supplier facing correlated, but stationary demand uses the sequence of forecast updates it receives to place orders upstream. Orders are constrained to be at most $C$ units. We assume that ordering cost is sufficiently high that if an order is placed in any given period, it will be a full-capacity order. The total cost of an order is denoted by $K$. The supplier also incurs inventory holding and backorder costs, denoted by $L(I_n)$, where $I_n$ is the end-of-period inventory level. At the beginning of period $n$,
demand $D_n$ is observed and forecasts are updated. Based on this information, the order quantity $P_n$ is determined, where $P_n \in \{0, C\}$. Orders are delivered after a delay of $l$ periods. Demands are satisfied using on-hand inventory and the newly received products such that the end-of-period inventory level $I_n$ evolves according to the relation $I_n = I_{n-1} + P_{n-l} - D_n$. The one-period cost in period $n$ is $K\delta(P_{n-l}) + L(I_n)$; this cost formulation assumes that ordering costs are incurred upon receipt of goods.

The supplier’s goal is to determine an ordering policy minimizing total expected cost, discounted at rate $\alpha$, over a finite horizon of $N$ periods. The terminal cost is a function $g(\cdot)$ of the ending inventory level. In the following formulation of the optimization problem, we assume that decisions are taken in periods 1 to $N$, and that costs are incurred in periods $l + 1$ to $l + N$ and discounted to period 1.

$$\min_{\{P_n, n=1,2,\ldots,N\}} E_{D_1,D_2,\ldots,D_{l+N}} \left[ \sum_{n=1}^{N} \alpha^{n+l}(K\delta(P_n) + L(I_{n+l})) + \alpha^{N+l}g(I_{N+l}) \right]$$

## 3 Analysis

We proceed by defining state variables whose evolution equations depend on iid disturbances and constructing a dynamic programming recursion based on these variables that solves the above optimization problem. Let $X_n$ denote the inventory on hand plus on order minus backorders at the beginning of period $n$, called the inventory position. Define $F_n = (F_{n,n+1}, F_{n,n+2}, \ldots, F_{n,n+H}, \lambda)$, the forecast vector recorded in period $n$ and consisting of forecasts for the next $H + 1$ periods. At the beginning of period 1, $F_0$ is available. The state variables $X_n$ and $F_n$ evolve according to

$$X_{n+1} = X_n + P_n - D_n$$

$$= X_n + P_n - (F_{n-1,n} + \varepsilon_{n,n}), \text{ and}$$

$$F_n = (F_{n-1,n+1} + \varepsilon_{n,n+1}, F_{n-1,n+2} + \varepsilon_{n,n+2}, \ldots, F_{n-1,n+H-1} + \varepsilon_{n,n+H-1}, \lambda + \varepsilon_{n,n+H}, \lambda).$$

Define $E_{n-1,n+i} = D_{n+i} - F_{n-1,n+i}$, the forecast error about period $n + i$ demand as of the end of period $n - 1$, and $E_{n-1}^k = \sum_{i=1}^{k} E_{n-1,n-1+i}$, the forecast error about the total demand over the next $k$ periods as of the end of period $n - 1$. Similarly,
define $F^k_{n-1} = \sum_{i=1}^{k} F_{n-1,n-1+i}$, the total forecast for the next $k$ periods as of the end of period $n - 1$. Then

$$I_{n+l} = X_n + P_n - \sum_{i=0}^{l} D_{n+i}$$

$$= X_n + P_n - \sum_{i=0}^{l} [F_{n-1,n+i} + (D_{n+i} - F_{n-1,n+i})]$$

$$= X_n + P_n - \sum_{i=0}^{l} [F_{n-1,n+i} + E_{n-1,n+i}]$$

$$= X_n + P_n - F_{n+1} - E_{n-1}.$$ 

In terms of the state vector $(X_n, F_{n-1})$, the one-period cost in period $n + l$ can be written as

$$K\delta(P_n) + L(X_n - F_{n-1}^{l+1} + P_n - E_{n-1}^{l+1})$$

and the optimization problem as

$$\min_{\{P_n, n=1,2,\ldots,N\}} E_{\varepsilon_1,\ldots,\varepsilon_N} \left[ \sum_{n=1}^{N} \alpha^{n+l}(K\delta(P_n) + L(X_n - F_{n-1}^{l+1} + P_n - E_{n-1}^{l+1})) \right.$$ 

$$+ \alpha^{N+l} g(X_{N+1} - F^{l}_{N} - E_{N}^{l})].$$

Let $J(X_1, F_0)$ denote the optimal cost of the optimization problem starting with an initial inventory position $X_1$ and forecast vector $F_0$. Define

$$\tilde{L}(y_n, F_{n-1}) = E_{\varepsilon_1,\varepsilon_{n+1},\ldots,\varepsilon_{N+1}} \tilde{L}(y_n - F_{n-1}^{l+1} - E_{n-1}^{l+1})$$

The dynamic programming recursion

$$f_{N+1}(X_{N+1}, F_N) = E_{\varepsilon_{N+1},\varepsilon_{N+2},\ldots,\varepsilon_{N+l}} [g(X_{N+1} - F_N - E_N^l)]$$

$$f_n(X_n, F_{n-1}) = \min \{ \tilde{L}(X_n, F_{n-1}) + \alpha E_{\varepsilon_n} [f_{n+1}(X_n - (F_{n-1,n} + \varepsilon_{nn}), F_n)],$$

$$K + \tilde{L}(X_n + C, F_{n-1}) + \alpha E_{\varepsilon_n} [f_{n+1}(X_n + C - (F_{n-1,n} + \varepsilon_{nn}), F_n)] \},$$

$$n = 1, \ldots, N$$

solves the optimization problem, that is, $f_1(X_1, F_0) = J(X_1, F_0)$. Note that, for brevity, we suppress the dependence of $F_n$ on $\varepsilon_n$ with respect to which expectation is taken. Define

$$J_n(y_n, F_{n-1}) = \tilde{L}(y_n, F_{n-1}) + \alpha E_{\varepsilon_n} [f_{n+1}(y_n - (F_{n-1,n} + \varepsilon_{nn}), F_n)]$$

5
and
\[ H_n(x_n, F_{n-1}) = K + J_n(x_n + C, F_{n-1}) - J_n(x_n, F_{n-1}). \]

Then
\[
\begin{align*}
f_n(X_n, F_{n-1}) &= \min\{J_n(X_n, F_{n-1}), K + J_n(X_n + C, F_{n-1})\} \\
&= J_n(X_n, F_{n-1}) + \min\{0, \bar{H}_n(X_n, F_{n-1}) - J_n(X_n, F_{n-1})\} \\
&= J_n(X_n, F_{n-1}) + \min\{0, H_n(X_n, F_{n-1})\}.
\end{align*}
\]

Define \( s_n(F_{n-1}) = \sup\{x \mid H_n(x, F_{n-1}) \leq 0\} \) with \( s_n(F_{n-1}) = -\infty \) if \( \{x \mid H_n(x, F_{n-1}) \leq 0\} = \emptyset \). The following result characterizes the optimal ordering region in period \( n \) in terms of the threshold \( s_n(F_{n-1}) \).

**Theorem** If \( f_{N+1}(x + C, z) - f_{N+1}(x, z) \) and \( \bar{L}(x + C, z) - \bar{L}(x, z) \) are both non-decreasing in \( x \), then there is an optimal policy of the form
\[
P_n(X_n, F_{n-1}) = \begin{cases} 
C & \text{if } X_n \leq s_n(F_{n-1}), \\
0 & \text{o/w}
\end{cases}
\]

\( n = 1, \ldots, N. \)

**Proof.** It is optimal to order in period \( n \) if \( H_n(X_n, F_{n-1}) \leq 0 \), that is, the ordering region in period \( n \) is given by the set \( \{(X_n, F_{n-1}) \mid H_n(X_n, F_{n-1}) \leq 0\} \). If \( H_n(x, z) \) is non-decreasing in \( x \), then the ordering region, parametrized by \( F_{n-1} \), is \( \{X_n \mid X_n \leq s_n(F_{n-1})\} \). We have that
\[
H_n(x, F_{n-1}) = K + J_N(x + C, F_{n-1}) - J_N(x, F_{n-1})
\]
\[
= K + \bar{L}(x + C, F_{n-1}) - \bar{L}(x, F_{n-1}) + \\
\alpha E_{\varepsilon_N}[f_{N+1}(x + C - (F_{n-1,N} + \varepsilon_{NN}), F_N) \\
- f_{N+1}(x - (F_{n-1,N} + \varepsilon_{NN}), F_N)]
\]

is non-decreasing in \( x \) due to the assumptions on \( f_{N+1} \) and \( \bar{L} \).

Suppose that \( H_{n+1}(x, F_n) \) is non-decreasing in \( x \) with \( s_{n+1}(F_n) = \sup\{x \mid H_{n+1}(x, F_n) \leq 0\} \). We will show that \( H_n(x, F_{n-1}) \) is nondecreasing in \( x \). Using \( f_{n+1}(x, F_n) = J_{n+1}(x, F_n) + \min\{0, H_{n+1}(x, F_n)\} \), we write
\[ K + f_{n+1}(x + C, F_n) - f_{n+1}(x, F_n) \]
\[ = K + J_{n+1}(x + C, F_n) - J_{n+1}(x, F_n) + \min \{0, H_{n+1}(x + C, F_n)\} - \min \{0, H_{n+1}(x, F_n)\} \]
\[ = H_{n+1}(x, F_n) + \min \{0, H_{n+1}(x + C, F_n)\} - \min \{0, H_{n+1}(x, F_n)\} \]
\[ = \max \{0, H_{n+1}(x, F_n)\} + \min \{0, H_{n+1}(x + C, F_n)\}. \]

Since both of these terms are nondecreasing in \(x\), we conclude that \(K + f_{n+1}(x + C, F_n) - f_{n+1}(x, F_n)\) is nondecreasing in \(x\). Consequently, \(E_{\varepsilon_n}[K + f_{n+1}(x - (F_{n-1,n} + \varepsilon_{nn}), F_n) - f_{n+1}(x - (F_{n-1,n} + \varepsilon_{nn}), F_n)]\) is nondecreasing in \(x\). Moreover, \(H_n(x, F_{n-1}) = K + J_n(x + C, F_{n-1}) - J_n(x, F_{n-1}) = K + \tilde{L}(x + C, F_{n-1}) - \tilde{L}(x, F_{n-1}) + \alpha E_{\varepsilon_n}[f_{n+1}(x - (F_{n-1,n} + \varepsilon_{nn}), F_n) - f_{n+1}(x - (F_{n-1,n} + \varepsilon_{nn}), F_n)].\) By the assumption on \(\tilde{L}\), we conclude that \(H_n(x, F_{n-1})\) is nondecreasing in \(x\). Thus, the ordering region in period \(n\), parameterized by \(F_{n-1}\), is given by \(\{X_n \mid X_n \leq s_n(F_{n-1})\}\), where \(s_n(F_{n-1}) = \sup \{x \mid H_n(x, F_{n-1}) \leq 0\}\).

As shown in Gallego and Toktay (2003), the class of functions satisfying the assumptions of the Theorem contains the class of functions \(h(x, z)\) that are absolutely continuous and convex in \(x\).

## 4 Conclusion

In this paper, we took the viewpoint of a supplier who receives exogenous forecast update information and wishes to determine an ordering policy to meet stationary stochastic demand at minimum total ordering, inventory holding and backorder cost. We used the Martingale Model of Forecast Evolution to characterize the sequence of forecast updates and imposed an all-or-nothing ordering structure on the supplier, an assumption justified in the presence of a large ordering cost. We formulated the dynamic programming recursion that solves the supplier’s problem as a function of a state variable vector consisting of the inventory position and the vector of forecasts over the forecast horizon. This formulation allowed us to prove that under non-restrictive assumptions on the one-period and end-of-horizon cost functions, the
optimal ordering policy is a state-dependent threshold policy with respect to the inventory position.

References


