Manufacturing & Service Operations Management

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

Gaining Benefits from Joint Forecasting and Replenishment Processes: The Case of Auto-Correlated Demand

Yossi Aviv,

Gaining Benefits from Joint Forecasting and Replenishment Processes: The Case of Auto-Correlated Demand

Yossi Aviv
Olin School of Business, Washington University, St. Louis, Missouri 63130
aviv@olin.wustl.edu

In this paper we consider a cooperative, two-level supply chain consisting of a retailer and a supplier. As in many practical settings, the supply chain members progressively observe market signals that enable them to explain future demand. The demand itself evolves according to an auto-regressive time series. We examine three types of supply chain configurations. In the first setting, the retailer and the supplier coordinate their policy parameters in an attempt to minimize systemwide costs, but they do not share their observations of market signals. In the second setting, resembling many original vendor-managed inventory (VMI) programs, the supplier takes the full responsibility of managing the supply chain's inventory, but the retailer's observations of market signals are not transferred to him. In our third setting, reminiscent of collaborative forecasting and replenishment partnerships, inventory is managed centrally, and all demand related information is shared. We propose a set of stylized models to study the three settings and use them to provide managerial insights into the value of information sharing, VMI, and collaborative forecasting.
(*Collaborative Forecasting; CPFR; Supply Chain Management; VMI.*)

1. Introduction

In an attempt to streamline their supply chains, companies have engaged in information sharing practices and have established various types of coordination mechanisms. In recent years, electronic data sharing has virtually become an industry standard, and has triggered several important practices such as *quick response (QR)* (Hammond 1990, Frazier 1986) and *efficient consumer response (ECR)* (Kurt Salomon Associates 1993). The initial benefits associated with electronic information sharing included the elimination of expensive paperwork and the reduction in human error. With the development of the *electronic data interchange (EDI),* companies have derived even larger benefits through lead time reduction (due to faster order processing), better visibility into inventory and *point-of-sale (POS)* data, and through better order tracking systems. For example, Wal-Mart uses EDI as well as the web-based *retail link* (which connects more than 4,000 suppliers) to allow suppliers to look at POS data to monitor the sales of their products. This, in turn, increases the level of service provided to Wal-Mart's customers due to the ability of the vendors to better plan and execute their business. In recent years, rapid evolution in information technology has paved the way for the implementation of a variety of coordination mechanisms in supply chains, and in many cases it has provided the necessary means for supply chain partners to restructure the way they conduct business. As an example, the reader can refer to Andersen Consulting (1998) for a study of the ways by which companies in the personal computer (PC) industry have integrated their supply chains; see also Magretta (1998) for a discussion of the specific case of Dell Computer.

Consider *vendor-managed inventory (VMI),* a supply

1523-4614/02/0401/0055$05.00
1526-5498 electronic ISSN

chain partnership that has received a lot of industry attention. In a VMI program, the supplier takes the responsibility for replenishing the customer's (e.g., retailer's) inventory. Practically, VMI transforms an originally decentralized inventory control setting into an inventory system controlled in a centralized fashion. To enable the supplier to perform the inventory management task under a VMI arrangement, POS data is usually transferred to the supplier by the retailer on a regular basis. It is often argued that VMI programs are beneficial, simply because they take a systemwide view of the entire chain. However, this assertion may be wrong if, for instance, POS data do not convey to the vendor all relevant information about future demand that exists in the system. For example, the retailer may have important information about an upcoming promotion, which is not conveyed via the history of sales data (in the food industry, knowledge about an upcoming salsa promotion should be taken into account by the corn chip supplier). If this kind of information is absent, it results in a vendor making replenishment decisions based on partial information. The lack of complete information may explain why a large number of VMI programs have failed because of poor supplier forecasting performance. The retailers involved in these failed partnerships often claimed that they ''could do a better job.'' For instance, Spartan Stores, Inc., a grocery chain, halted its VMI programs after 12 months of operations, blaming part of the failure on the fact that the VMI vendors had not taken promotions data into account (see Mathews 1995). Furthermore, K-Mart cut VMI contracts from more than 300 to about 50 (Fiddis 1997), blaming poor performance (partially) on the fact that the suppliers did not have adequate forecasting skills. Therefore, we conclude that it is important to integrate the dimension of forecasting capabilities into the study of VMI partnerships.

The fact that centralized settings may perform poorly if local knowledge cannot be shared with the decision maker is not conceptually new; see Hayek (1945) and, more recently, Anand and Mendelson (1997). However, it has taken the industry several decades before certain supply chains have become extensively integrated, by coordinating key business processes and by sharing critical knowledge throughout their channels. Particularly, the Voluntary Interindustry Commerce Standard (VICS) Association established the *collaborative planning, forecasting and replenishment (CPFR)* subcommittee in 1993, with the aim of developing standards for collaborative processes for the industry (Koloczyc 1998, Aviv 2001a). In our conversations with managers that have been directly involved with CPFR, the latter have claimed that although companies embrace the CPFR concept enthusiastically, they often lack the means to assess its benefits in their own settings. To date, several CPFR pilot studies have provided promising success stories, but their implementations have been limited in scale, and hence companies are still hesitant in embarking on such initiatives on a larger scale.

To explore the forecasting issues involved in the implementation of VMI and CPFR programs, we have decided to construct a stylized framework that would capture the following aspects: First, our interest is focused on cooperative supply chains. In other words, we assume that the supply chain members are concerned only with the overall supply chain cost. Second, our models need to be rich enough to capture the forecasting-related knowledge that exists in the system. To this end, we incorporate early demand information in the form of explanatory variables through which the supplier and the retailer learn about future demand (a similar approach was taken in the recent paper of Aviv 2001a). In addition, the literature suggests that the traditional, i.i.d. demand assumption may not be appropriate to describe practical settings because demands in different periods are often correlated; see, e.g., Kahn (1987), Graves (1999), and Lee et al. (2000). Hence, we make use of the AR(1) time-series model in this work. Finally, the interaction between the trading partners' information sets needs to be modeled. To achieve this goal, we examine three types of two-stage supply chain configurations, each consisting of a single retailer and a single supplier. In the first setting, the retailer and the supplier coordinate their policy parameters in an attempt to minimize systemwide costs, but they do not share their observations of explanatory market signals. We name this model *locally managed inventory*

(LMI) because inventory is controlled at each level using local information only. In our second setting, named *supplier-managed inventory (SMI),* the supplier takes the full responsibility of managing the supply chain's inventory, but the retailer's observations of market signals are not transfered to him. Therefore, this model resembles many original VMI programs, in which forecasting data available to the retailer is not fully shared with the vendor. The SMI setting is typical to cases in which the retailer and the supplier do not have the required infrastructure, such as communication links between inventory planning systems and statistical software packages, to support a smart exchange of forecast information (hence, SMI should not be viewed as a VMI setting in which the supplier naively ''throws out'' information provided by the retailer). In our third setting, reminiscent of CPFR partnerships, inventory is managed centrally, and all demand related information is shared. We name this third setting *collaborative forecasting and replenishment (CFAR).* In all of our settings, we assume that the supply chain members have a complete knowledge of the underlying system parameters, such as cost rates, and the characteristics of the forecasting and demand processes. We deliberately use the terms SMI and CFAR instead of VMI and CPFR (respectively) because there are certain important aspects of the latter programs that are not modeled in this paper and because, in practice, VMI does not necessarily imply that no forecasting information is shared between the retailer and the supplier.

Intuitively, CFAR is expected to have the best inventory performance among the three settings. Yet, CFAR may require substantial investment, compared to LMI and SMI settings: CFAR usually entails a strong commitment of organizational resources and a willingness to share not only POS data but other types of information, such as promotion plans, early placement of customer orders, and perhaps even a history of sales of competing products. For instance, Lee and Whang (2001) assert that to date only a relatively small number of companies have been successful in using rich sets of well chosen and timely data to drive their replenishment processes. It is therefore our objective to provide some general guidelines as to when each one of the settings, SMI or CFAR, is expected to bring substantial inventory-related gains to the supply chain. We discuss this subject (see, e.g., §6.3), through a study of the *magnitude* of the differences in cost performance between the three settings.

The management science literature has paid a great deal of attention to the area of supply chain coordination in the past several years. The topic of information sharing in supply chains is discussed in Aviv and Federgruen (1999), Cachon and Fisher (2000), Chen (1998), Gavirneni et al. (1999), Lee et al. (2000), and Lee and Whang (1999); rather than duplicate these papers, we refer the reader to their discussions of the literature. Aviv and Federgruen (1999) study the operational benefits of VMI programs in a single-supplier, multiretailer supply chain structure, with capacity constraints and fixed ordering costs. Cheung and Lee (2002) also examine the benefits of VMI programs in multiple retailer settings. They identify two main advantages for VMI: the ability to coordinate shipments to retailers and the ability to rebalance the stocking positions of the various retailers. These papers assume that the VMI supplier receives full POS data and knows the inventory position across the entire chain at any point in time. Yet, in their models, no early information about future consumer demand can be obtained by any member of the supply chain.

In recent years, researchers have suggested several stylized models to study the problem of inventory control when demand information is not commonly known throughout the supply chain. Chen et al. (2000a) propose a two-stage supply chain model in which neither the retailer nor the supplier know the exact form of the demand process, which is modeled as an auto-regressive process of order 1. They assume that at the beginning of each ordering period, the retailer uses a simple moving average procedure to generate two values (mimicking a mean and a standard deviation), based on which orders are determined. Chen et al. then investigate the magnitude of the so-called ''*bullwhip phenomenon,*'' or the level by which the variability of order quantities is higher than the variability of the actual demand. Watson and Zheng (2000) analyze a supply chain facing a serially correlated demand model. Here, again, the supply chain partners do not know the exact characteristics of the demand process. The authors claim that inventory managers tend to overreact to demand changes and show that this in turn leads to suboptimal replenishment policies internally and causes an undue bullwhip effect externally. Croson and Donohue (2001) conduct an experimental study, based on the popular Beer Distribution Game, to uncover behavioral factors that impact a company's ability to manage its supply chain. They measure how the institutional designs of supply chains impact the magnitude and timing of the bullwhip effect in a multiechelon system. For other works investigating the impact of forecasting in cases where the underlying characteristics of the demand distribution are unknown to the decision makers, see, e.g., Miyaoka and Kopczak (2000) and Chen et al. (2000b).

Another approach to the study of supply chains in which demand information is not commonly known was taken recently by Aviv (2001a). The main difference between his work and the papers described above is that Aviv assumes that the supply chain members are able to fully characterize the demand process in a linear regression form that depends on early market signals. However, these signals are observed locally by the supply chain partners, so forecasting and inventory management may be based on partial information only. Aviv then discusses the value of comanaged forecasting processes in *decentralized* supply chains, where he shows that joint forecasting processes may provide substantial benefits to supply chains. Nevertheless, when supply chain members do not have unique forecasting skills and when the level of information available to each is not substantially diversified, comanaged forecasting processes tend to be risky. Our paper takes a similar approach to that of Aviv (2001a), yet it differs in several ways. First, Aviv assumes that demands in different periods are independent, while we deal with an auto-regressive demand process. Recently, Raghunathan (2001) has specifically suggested that a model of auto-regressive demand, like the one discussed in our paper, would be useful for explaining the value of information sharing in supply chains. Second, Aviv (2001a) has not focused specifically on the choice between VMI and CPFR programs, two important supply chain partnerships. Finally, the supplier's forecasting process in Aviv (2001a) used approximations for the best (minimum mean square) estimates of future demand. Here, using a recent framework by Aviv (2001b), we develop the *optimal* estimates for all of our settings.

The remainder of this paper is organized as follows: §2 provides the basic notation and preliminaries for this paper. Section 3 describes the inventory replenishment processes in the three supply chains. In §4 we develop the forecasting procedures for all settings, and in §5 we propose a method for calculating the inventory control policies that minimize the long-run average cost for the supply chain. In §6 we discuss our numerical study and provide some managerial insights. Section 7 concludes this paper.

2. Notation and Preliminaries

We focus on a simple supply chain consisting of a single retailer and a single supplier. The retailer's and the supplier's inventories are replenished periodically. The supplier orders from an unlimited source of supply. Orders placed by the supplier are received after a lead time of τ periods, and any delivery from the supplier to the retailer, made at the beginning of a particular period, is received L periods later. We assume that shortages are fully backlogged. At the end of each period, the supply chain is charged according to the following cost structure: (i) at the retailer's level—a holding cost at a rate of h^r per unit of inventory on hand, plus a shortage penalty cost at a rate of p^r per unit of backordered demand and (ii) at the supplier's level—a holding cost at a rate of h^s per each unit of inventory held from the beginning to the end of the period.

2.1. The Demand Model

We model the demand as an auto-regressive statistical time series of order 1, as follows:

The underlying AR(1) model:

$$d_t - \mu = \alpha(d_{t-1} - \mu) + q_t, \qquad (1)$$

where μ represents the long-run average demand per period, and α is a known constant ($0 \leq \alpha < 1$). The sequence $\{q_t\}$ consists of independent and identically distributed (i.i.d.) random components, each of which is normally distributed with a mean of zero. The random components $\{q_t\}$ represent a sequence of "shocks," where q_t has a full impact on the demand during period t, and a partial, diminishing impact on future demands $(d_{t+1}, d_{t+2}, \ldots)$.

This model for demand is common for describing settings in which intertemporal correlation between demands in consecutive periods exists (see, e.g., Chen et al. 2000a, Lee et al. 2000, and Kahn 1987). We now propose an extension to model (1) that will enable us to accommodate situations in which supply chain members are able to learn about future demand through a variety of early *signals*, as described in the Introduction. Specifically, suppose that the random error term q_t can be *explained* statistically by the general linear regression model

$$q_t = \theta_r \sum_{i=1}^{T} \delta_{t,i}^r + \theta_s \sum_{i=1}^{T} \delta_{t,i}^s + \epsilon_t, \qquad (2)$$

where the $\delta_{t,i}^r$ and $\delta_{t,i}^s$, can be fully determined by the retailer and the supplier, respectively, during period $t - i$. For example, $\delta_{t,i}^r$ by itself can be a function of new information about the expected weather condition in period t, early order placement by customers, etc. Hence, the δ variables in (2) represent a set of locally observable *explanatory variables*, and ϵ_t is an independent, unobservable error component. We refer to $\delta_{t,i}^m$ as the *market signal* about q_t observed during period $t - i$, by member m of the supply chain ($m \in \{r, s\}$ represents the retailer and the supplier, respectively). Specifically, we assume that the market signals are observed by the supply chain members *progressively* over time, beginning from a maximal horizon of T periods in advance. We further assume that $\delta_{t,i}^m \sim N(0, \sigma_{m,i})$ and $\epsilon_t \sim N(0, \sigma_0)$. Clearly, the information that the retailer and the supplier have about future demand are expected, in general, to be correlated. Therefore, while we assume that the δ and ϵ components are independent, we make the important exception that for every combination of $t \geq 1$ and $i = 1, \ldots, T$, the couple $(\delta_{t,i}^r, \delta_{t,i}^s)$ can be correlated. To this end, we assume a fixed correlation that is independent of t and i, and we denote this correlation by the parameter ρ. Hence, it follows that the variance of q_t is equal to:

$$\mathrm{Var}(q_t) = \sigma_0^2 + \sum_{i=1}^{T} [\theta_r^2\sigma_{r,i}^2 + 2\rho\theta_r\theta_s\sigma_{r,i}\sigma_{s,i} + \theta_s^2\sigma_{s,i}^2]. \qquad (3)$$

In the linear regression model, the parameters θ_r and θ_s represent, to a certain degree, the explanatory strength of the retailer and the supplier. But to sharpen this interpretation, it is useful to examine the possible reduction in the uncertainty about q_t that can be achieved by observing the sequence $\{\delta_{t,i}^m\}$. Because of the dependency between $\delta_{t,i}^r$ and $\delta_{t,i}^s$, the possible *reduction* in the unexplained variability of q_t during period $t - i$ is: $(\theta_r\sigma_{r,i} + \rho\theta_s\sigma_{s,i})^2$ for the retailer and $(\rho\theta_r\sigma_{r,i} + \theta_s\sigma_{s,i})^2$ for the supplier. We interpret the ratio between the two latter values as the *relative explanation power* of the retailer; see §6.3 for further discussion. For instance, if $\rho = 0$ and $\sigma_{r,i} = \sigma_{s,i}$, the meaning of $\theta_r > \theta_s$ is that the retailer's market signals are more informative than the supplier's market signals. As we mentioned above, we assume that in all three of our settings the retailer and the supplier know the *parameters* of the model (1)–(2).

3. The Inventory Replenishment Processes

In this section we describe the inventory replenishment policies for each of our three settings. Our main focus is on minimizing the long-run average total supply chain cost per period; hence we assume that in each setting the retailer's and supplier's replenishment policy parameters will be set in an attempt to achieve this goal. The following sequence of events, within every period t, applies:

(i) Forecasts are adjusted, based on new data observed during the previous period. In the LMI setting, each member (m) makes the forecast adjustment separately. Also, in this setting, the supplier makes the forecast adjustment *after* observing the retailer's order for the specific period. In the CFAR setting, there exists a single, joint forecasting system that is based on the retailer's and supplier's observations (δ variables). In the SMI, only the supplier's observations are used.

(ii) Replenishment orders and deliveries of goods are made. In the LMI setting, we assume that the retailer's orders are placed with the supplier, before the latter places his orders. In the SMI setting, the supplier is responsible for the replenishment process at both levels. In the CFAR setting, a centralized decision maker (e.g., the supplier, the retailer, or a third party) is responsible for replenishing the inventories at both levels.

(iii) Demand is realized, and inventory holding costs and shortage penalties are charged.

Ideally, one would like to identify an optimal inventory policy for each setting, and to be able to evaluate the long-run average costs under such an optimal policy. However, the non-i.i.d. demand pattern, combined with the forecasting evolution process described above, makes such a task prohibitive. At best, we expect that the dynamic program of each setting will yield an optimal policy that prescribes *state-dependent* order quantities (see, e.g., Song and Zipkin 1993), where the multidimensional state variable includes the inventory positions at both levels (or echelons), as well as sufficient statistics for the history of the explanatory variables. Apparently, the high dimensionality of the state space makes the construction of optimal policies intractable and, more importantly, impractical. Our approach is therefore to restrict ourselves to a more convenient and intuitively appealing class of policies, and it is with respect to such a policy class that we compare our three settings. The main guiding principle of our restricted policy class is that for each level of the supply chain, two values are taken into account:

(i) the inventory position of the level (local or echelon position, depending on the particular settings) and

(ii) the best estimate of total future orders that will be placed during a lead time. We shall be more specific later about which demand and which lead times are used in each case.

In all settings, orders for each installation/echelon are placed according to: an adaptive, minimum mean-square error (MMSE) base-stock policy—*order enough units to bring the inventory position as close as possible to your best (in the sense of MMSE) estimate of future lead-time demand, plus a predetermined, fixed "safety stock."*

Clearly, because the forecasts are uniquely determined (MMSE), it is the choice of the safety stock quantities that fully defines the overall inventory replenishment process within our restricted class of policies. Therefore, we need to deal with two tasks in all of our settings: First, in §4, we describe the way by which MMSE forecasts of lead-time demands are calculated for each level of the supply chain, at the beginning of every period. Second, we show in §5 how the levels of safety stock are set in a *coordinated* fashion, in search of minimum long-run average total supply chain costs.

4. The Forecasting Processes

In this section, we develop forecasting procedures for our LMI setting (§4.2) and for our SMI and CFAR settings (§4.3). The forecasting processes we develop below provide the means to compute the MMSE estimates of future lead-time demand at the retailer's and supplier's levels, so that they can be used in determining the period-by-period inventory replenishment quantities. This task, however, is not simple. Consider, for instance, the supplier's objective of forecasting the volumes of the next several orders to be placed by the retailer in the LMI setting. To this end, the supplier would need to use his own accumulated information (market signals) about future demand, as well as the *history* of the retailer's orders. However, because of the nontrivial correlation between these two streams of information, the statistical task is quite tedious. Fortunately, we are able to show that the model (1)–(2) maintains a linear state space form, for which Aviv (2001b) has recently constructed a unified framework for forecasting and inventory management. Aviv (2001b) considers a locally managed, two-stage supply chain that faces a demand process that maintains a so-called linear state space form. His framework can take into account situations in which the supply chain members partially or fully observe information about future demand, and is hence useful in our analysis. To keep this paper self-contained we briefly describe in the following subsection what we mean by a linear state space form.

4.1. Preliminaries

A key observation for our analysis is that the demand process can be represented in a *linear state space form.* In other words, suppose there exists a state vector X_t, which follows the evolution $X_t = FX_{t-1} + V_t$, for a given *known* square matrix F, and a random vector process $\{V_t\}$ that consists of components that are independent across time, each having a fixed (time-invariant) multivariate normal distribution with a mean vector of zero. We denote the covariance matrix of V_t by Σ_V. The demand d_t is assumed to be given by $d_t = GX_t + \mu$, for a known vector G. In the context of this paper, the vector X_t may include sufficient statistics of all market signals and demand realizations. For such linear state space form, one can develop forecasting procedures of relatively simple recursive structure, utilizing the *Kalman filter technique.* Aviv (2001b) considers the following situation: Suppose that a forecaster observes by the end of every period t only part of the information embedded in X_t. Specifically, for a given matrix H, assume that the forecaster fully observes the linearly filtered vector $\Psi_t = HX_t$. Next, define $\hat{X}_{t\,|\,t-1} = \mathrm{E}[X_t \,|\, \Psi_{t-1}, \Psi_{t-2}, \ldots]$ to be the MMSE estimate of X_t, generated by the forecaster at the beginning of period t, based on the history of his observations $\{\Psi_{t-1}, \Psi_{t-2}, \ldots\}$. Let $\Omega^X_{t\,|\,t-1} \doteq \mathrm{E}[(X_t - \hat{X}_{t\,|\,t-1})(X_t - \hat{X}_{t\,|\,t-1})']$ be the associated error covariance matrix. The following recursive scheme[1] is used as the basis for the forecasting process:

$$\hat{X}_{t|t-1} = F\hat{X}_{t-1|t-2} + FK_{t-1}(\Psi_{t-1} - H\hat{X}_{t-1|t-2}), \quad (4)$$

$$-^{X}_{t|t-1} = F -^{X}_{t-1|t-2}F' - FK_{t-1}H\,\Omega^X_{t-1|t-2}F' + \Sigma_V, \quad (5)$$

where $K_{t-1} = \Omega^X_{t-1\,|\,t-2}H'(H\,\Omega^X_{t-1\,|\,t-2}H')^{-1}$. Because we are interested in long-run performance in this work, we shall focus our attention on the Kalman filter in *steady state,* and hence ignore the subscripts in the matrices $-^X$ and K hereforth. Aviv (2001b) showed that the conditional distribution of the estimate of the aggregate demand during periods $t, \ldots, t + L$, given the information available to the forecaster, is normal with mean

$$(L + 1)\mu + G\left(\sum_{l=0}^{L} F^l\right)\hat{X}_{t|t-1}, \quad (6)$$

and variance

$$\sum_{l=0}^{L-1}\left[G\left(\sum_{j=0}^{l} F^j\right)\Sigma_V\left(\sum_{j=0}^{l} F^j\right)'G'\right] + G\left(\sum_{j=0}^{L} F^j\right)\Omega^X\left(\sum_{j=0}^{L} F^j\right)'G'. \quad (7)$$

4.2. Forecasting in the Locally Managed Inventory Setting

The LMI setting represents a situation in which each member of the supply chain replenishes his own inventory by following an *installation-based,*[2] adaptive base-stock policy. Under such a policy, each member considers his local (i.e., installation-based) inventory position, defined as the number of units on hand, minus backlogs (at his particular location), plus all outstanding orders. The member then computes his best estimate of an associated lead-time demand. For the retailer, the lead-time demand is given by $d_t + \cdots + d_{t+L}$, whereas for the supplier it is given by $A_{t+1} + \cdots + A_{t+\tau}$ (where A_t denotes the retailer's order at the beginning of period t). Note that A_t need not be estimated by the supplier during period t because the retailer places his order before the supplier needs to place his. Denote by $\hat{D}^{(L)}_t$ and $\hat{A}^{(\tau)}_t$ the MMSE forecasts of these two lead-time demand quantities, respectively. Recall that each of these estimates is based on *local* information only. Then, for each setting, our policy prescribes order quantities that are based on the two base-stock levels:

$$\text{Retailer's base-stock level:} \qquad \beta^r_t = \hat{D}^{(L)}_t + \gamma^r,$$

$$\text{Supplier's base-stock level:} \qquad \beta^s_t = \hat{A}^{(\tau)}_t + \gamma^s. \quad (8)$$

Next, we describe the way by which MMSE forecasts are generated and how the uncertainty surrounding them can be calculated. To this end, we identify for the LMI setting a linear state space form of the type described in the previous section. Specifically, we shall use the following $(2T + 1)$-dimensional state

[1]This is essentially the Kalman filter algorithm.

[2]For a discussion of installation and echelon policies, see Axsäter and Rosling (1993).

vector, denoted below by X_t (with X_t' denoting X_t transposed):

$$X_t' = (d_t - \mu, \Delta^r_{t+1,1}, \Delta^r_{t+2,2}, \ldots, \Delta^r_{t+T,T}, \Delta^s_{t+1,1}, \Delta^s_{t+2,2}, \ldots, \Delta^s_{t+T,T}).$$

Particularly, the information included in X_t is:

(i) The difference $d_t - \mu$, which is based on the most recent demand observation (this is sufficient information for predicting future demand in the standard AR(1) demand process; see Lee et al. 2000).

(ii) $\Delta^m_{t,l} \doteq \Sigma^T_{i=l}\, \delta^m_{t,i}$ ($l = 1, \ldots, T$, $m = r, s$). Each variable $\Delta^m_{t+l,l}$ represents the accumulated information about q_{t+l} acquired by member m up until the end of period t.

4.2.1. The Retailer's Forecasting Problem. To devise a forecasting process for the retailer, observe that the state space follows the evolution $X_t = FX_{t-1} + V_t$ where

$$F_{(2T+1)\times(2T+1)} = \begin{pmatrix} \alpha & \theta_r & 0 & \cdots & 0 & \theta_s & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & 0 & \vdots & & & \vdots \\ \vdots & \vdots & & \ddots & 1 & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & & & \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & & & \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

$$V_t = \begin{pmatrix} \epsilon_t \\ \delta^r_{t+1,1} \\ \vdots \\ \delta^r_{t+T,T} \\ \delta^s_{t+1,1} \\ \vdots \\ \delta^s_{t+T,T} \end{pmatrix},$$

and that $\Psi^r_t = HX_t$ is the $(T + 1)$-dimensional vector representing the information about X_t actually observed by the retailer. Clearly, Ψ^r_t includes the demand realization observed during period t, and all Δ^r data, and so $H = (I_{(T+1)\times(T+1)}\ \mathbf{0}_{(T+1)\times T})$. Furthermore, the demand equation is given by $d_t = GX_t + \mu$ where $G = (1, 0, 0, \ldots, 0) \in \mathbf{R}^{2T+1}$. Hence, the retailer's estimate of the lead-time demand, $\hat{D}^{(L)}_t$, is given by (6), where the estimates $\hat{X}_{t|t-1}$ can be calculated using the Kalman filters described in §4.1, with Ψ^r replacing Ψ. The mean-square error associated with the estimate $\hat{D}^{(L)}_t$ can be computed via (7). While the above state representation X_t is intuitively clear and encompasses the aggregate information available in the system, it turns out that one can achieve further reduction in the dimension of X_t if one is interested only in the retailer's forecasts. For the supplier, however, a similar reduction would not be feasible, as we shall argue later. The following proposition describes the retailer's forecasting process in a simplified way:

Proposition 1 (Retailer's Forecasts; LMI). *A further simplification of the estimates $\hat{D}^{(L)}_t$ can be obtained as follows: Let $\hat{\delta}^r_{t,i} \doteq (\theta_r + \theta_s\rho\sigma_{s,i}/\sigma_{r,i}) \cdot \delta^r_{t,i}$, and $\hat{\Delta}^r_{t,i} = \Sigma^T_{l=i}\, \hat{\delta}^r_{t,l}$. Also, define $G_* = (1, 0, \ldots, 0)$, $\Sigma_* = diag(\sigma_0^2 + (1 - \rho^2)\theta_s^2 \Sigma^T_{i=1} \sigma^2_{s,i}, (\theta_r\sigma_{r,1} + \rho\theta_s\sigma_{s,1})^2, \ldots, (\theta_r\sigma_{r,T} + \rho\theta_s\sigma_{s,T})^2)$, and*

$$F_* = \begin{pmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & & \ddots & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Then,

$$\hat{D}^{(L)}_t = (L + 1)\mu + \left(\sum_{l=1}^{L+1} \alpha^l\right)\cdot(d_{t-1} - \mu) + \sum_{j=1}^{\min(T,L+1)} \left(\sum_{l=0}^{L+1-j} \alpha^l\right)\cdot \hat{\Delta}^r_{t+j-1,j}, \tag{9}$$

$$\Omega^{LMI}_D = \sum_{l=0}^{L} \left\{ G_*\left(\sum_{j=0}^{l} F^j_*\right)\Sigma_*\left(\sum_{j=0}^{l} F^j_*\right)' G_*' \right\}. \tag{10}$$

(All proofs appear in the Appendix.)

The variable $\hat{\delta}^r_{t,i}$ in this proposition can be interpreted as the expected value of the conditional distribution of $(\theta_r\delta^r_{t,i} + \theta_{s}\delta^s_{t,i})$, given the value $\delta^r_{t,i}$. In other words, it is the best update to the estimate of q_t that the retailer can provide on the basis of the market signals he has observed during period $t - i$. The first two terms in (9) are precisely the same as we expect from the forecast expression for the basic AR(1) model (see, e.g., Lee et al. 2000). The third part of (9) incorporates all relevant market signals observed by the retailer up to time t. The reason for multiplying the signal $\hat{\Delta}^r_{t+j-1,j}$ by $\Sigma_{l=0}^{L+1-j}\alpha^l$ is that this signal does not only affect the demand d_{t+j-1} but also future demands, albeit by a diminishing rate α.

4.2.2. The Supplier's Forecasting Process. To suggest a forecasting process for the supplier, we first need to understand the way by which the retailer's orders are generated. Recall from (8) that the retailer uses the base-stock level $\hat{D}_t^{(L)} + \gamma^r$, and thus the order quantity generated at the beginning of period t is *approximately* given by

$$A_t = \hat{D}_t^{(L)} - \hat{D}_{t-1}^{(L)} + d_{t-1}, \tag{11}$$

as long as the values of $\hat{D}_t^{(L)} - \hat{D}_{t-1}^{(L)} + d_{t-1}$ are rarely negative. Because Equation (11) is quite appealing,[3] we make two assumptions below which will allow us to comfortably use it. First, suppose that the (unconditional) standard deviation of the error terms q_t is small relative to the mean demand μ (say $\mathrm{Std}(q_t)/\mu \le 0.25$). Aviv (2001a) showed that under this assumption, and when $\alpha = 0$, (11) is a good approximation for the *actual* order quantities. Second, Lee et al. (2000) argue (see Lemma 1 there) that:

Lemma 1. *For $\alpha \ge 0$, $P(A_t \ge 0)$ is increasing in α as $t \to \infty$.*

(Recall that in our analysis we are interested in long-run performance.) Thus, we shall limit our discussion to cases in which $\alpha \ge 0$ and $\mathrm{Std}(q_t)/\mu$ is small.

When estimating the lead-time demand $\Sigma_{l=1}^{\tau} A_{t+l}$, the supplier needs to take into account both the history of orders placed by the retailer (i.e., $\{A_t, A_{t-1}, \ldots\}$) and the history of market signals he has observed. To simplify the exposition of the supplier's forecasting procedure, we consider the case in which $\{\sigma_{r,i} = c\cdot\sigma_{s,i}: i = 1, \ldots, T\}$ for a given constant c, and without loss of generality, we shall assume $c = 1$. This condition is assumed in our numerical study. The purpose of the next proposition is to show that from the supplier's perspective, the retailer's order in period $t + 1$ is a linear function of a certain state vector Z_t, and that the supplier observes a linear filtration of this state during period t. Furthermore, the proposition demonstrates that A_{t+1}, Z_t, and the observed state follow the linear state space described in §4.1.

Proposition 2. *Suppose that $\sigma_{r,i} = \sigma_{s,i}$ for all $i \ge 1$, and that the retailer's forecasting system is in steady state. Define the state vector $Z_t \doteq (HX'_{t-1}\ \ X'_t)'$. Then, the following system represents the coevolution of the retailer's order quantities and the supplier's observations:*

$$\underbrace{\begin{pmatrix} HX_{t-1} \\ X_t \end{pmatrix}}_{Z_t} = \underbrace{\begin{pmatrix} \mathbf{0}_{(T+1)\times(T+1)} & H \\ \mathbf{0}_{(2T+1)\times(T+1)} & F \end{pmatrix}}_{\tilde{F}} \cdot \underbrace{\begin{pmatrix} HX_{t-2} \\ X_{t-1} \end{pmatrix}}_{Z_{t-1}} + \underbrace{\begin{pmatrix} \mathbf{0}_{(T+1)\times 1} \\ V_t \end{pmatrix}}_{\tilde{V}_t}$$

$$\Psi_t^s = \begin{pmatrix} A_{t+1} \\ QX_t \end{pmatrix} = \underbrace{\begin{pmatrix} -M & G + MH \\ \mathbf{0}_{T\times(T+1)} & Q \end{pmatrix}}_{\tilde{H}} \cdot Z_t,$$

where $Q = (\mathbf{O}_{T\times(T+1)}\ \ I_{T\times T})$, and the vector $M \in \mathbf{R}^{T+1}$ is given by:

$$M_i = \begin{cases} \sum_{l=1}^{L+1} \alpha^l & i = 0, \\ (\theta_r + \rho\theta_s) \sum_{l=0}^{L+1-j} \alpha^l & 1 \le j \le \min(L+1, T), \\ 0 & \textit{otherwise}. \end{cases}$$

with the index j ranging between 0 and T. Finally, note that the process $\{\tilde{V}_t\}$ consists of independent components, and that

$$A_{t+1} = \underbrace{\tilde{R}\tilde{H}}_{\tilde{G}} Z_t; \qquad \tilde{R} = (1, 0, \ldots, 0).$$

It is instructive to note that the supplier's observable vector Ψ_t^s includes, in addition to the market signals observed by the supplier (i.e., the QX_t component), the retailer's order[4] A_{t+1}. This is essential for the forecasting process, because the supplier has to consider both types of information when generating

[3]And it is assumed extensively in the inventory management literature on information sharing in supply chains.

[4]We refer to A_{t+1} rather than A_t, because the retailer places his order before the supplier needs to place his. As mentioned, Ψ_t^s is observed after the beginning of period t but before the supplier's decision time in period $t + 1$.

forecasts of future orders. After all, the history of orders placed by the retailer provides some information about the latter's future orders. While it is hard to explain the intuition behind each of the various matrices defined in Proposition 2, it is perhaps more worthy to focus on the overall contribution of this proposition: Using the above linear state space representation, the supplier can calculate the estimates of future lead-time demands using the exact same mechanism as is used in the retailer's forecasting process. Here, the values of $\hat{Z}_{t|t-1}$ and the error covariance matrices $\Omega^Z_{t|t-1}$ are generated through the Kalman filter algorithm (4)–(5), merely replacing F, H, Ψ, and Σ_V with $\tilde{F}$, $\tilde{H}$, Ψ^s, and $\Sigma_{\tilde{V}}$, respectively. We conclude that, analogously to the retailer's forecasting process, the supplier's forecasts can be calculated as follows: At the beginning of period t, immediately after observing A_t, the supplier's estimate of the aggregate order quantity $\Sigma_{l=1}^{\tau} A_{t+l}$ is given by $\hat{A}_t^{(\tau)} = \tau\mu + \tilde{G}(\Sigma_{l=0}^{\tau-1} \tilde{F}^l) \cdot \hat{Z}_{t|t-1}$. In the long run, the mean-square error of this estimate is given by

$$\Omega_A^{\mathrm{LMI}} = \sum_{l=0}^{\tau-2} \left[\tilde{G}\left(\sum_{j=0}^{l} \tilde{F}^j\right)\Sigma_{\tilde{V}}\left(\sum_{j=0}^{l} \tilde{F}^j\right)' \tilde{G}' \right] + \tilde{G}\left(\sum_{j=0}^{\tau-1} \tilde{F}^j\right)\Omega^Z\left(\sum_{j=0}^{\tau-1} \tilde{F}^j\right)' \tilde{G}',$$

where Ω^Z is the steady state error covariance matrix of the supplier's estimation procedure.

REMARK. Our assumption that $\sigma_{r,i} = c \cdot \sigma_{s,i}$ for all i, and for a constant c, was made to ease the exposition of this section. Yet, the more general case of the arbitrary $(\sigma_{r,i}, \sigma_{s,i})$ pattern is still treatable using our framework: While the order quantities $\{A_t\}$ will keep evolving according to the linear state space model, one would have to modify the state vector Z_t, which will then have a higher dimensionality than at present $(3T + 2)$. Such possible extension is desired, for instance, if one wishes to model settings in which the retailer's market signals are more useful for predicting demands that will occur in the near future, while the supplier's market signals are more useful for predicting demands in the more-distant future.

4.3. Echelon-Based Inventory Systems

In this subsection we consider our echelon-based settings, SMI and CFAR. We assume that in both of these settings there is a single decision maker (the *supplier,* for the purpose of our discussion) that has the responsibility of replenishing the inventories both at the retailer's site and at the supplier's site. To make such a task practical, the supplier would need to gain access to inventory status information at the retailer's level. With this information available, the supplier is able to determine the replenishment quantities for the retailer by using the same type of adaptive MMSE base-stock policy that the retailer uses in the locally managed settings. However, here the supplier would be in charge of forecasting the demand $d_t + \cdots + d_{t+L}$ to be used in the determination of the base-stock level for the retailer's inventory position. For the supplier itself, the inventory management literature (e.g., Clark and Scarf 1960) suggests that an *echelon-based* policy should replace the supplier's installation-based policy discussed above.[5] Under an echelon-based policy, the supplier looks at the inventory position of the *entire* supply chain, instead of his local inventory position. This echelon inventory position consists of all goods within the supply chain (on hand, plus in-transit from the supplier to the retailer), minus backlogs at the retailer level (only), plus all outstanding orders at the supplier's level. The echelon position is then compared to the supplier's best estimate of the total demand during a lead time of $L + \tau + 1$ periods $(d_t + \cdots + d_{t+L+\tau})$, plus a fixed safety stock. To summarize, for each echelon-based setting our policy prescribes order quantities that are based on the two base-stock levels:

Retailer's echelon base-stock level:

$$B_t^r = \hat{D}_t^{(L)} + \Gamma^r,$$

Supplier's echelon base-stock level:

$$B_t^s = \hat{D}_t^{(L+\tau)} + \Gamma^s, \tag{12}$$

and the overall inventory replenishment process is hence fully defined by the choice of the safety stock parameters Γ^r and Γ^s.

The treatment of the SMI and CFAR settings is easier than that of the LMI setting, because decisions are

[5] Nevertheless, we show later that under the assumptions of our model, the echelon-based policies described in this paper can be easily converted into installation-based policies.

based on centralized information (even if the information is incomplete, such as in the SMI case). Essentially, the linear model (2) used for explaining the random errors $\{q_t\}$ based on observable market signals can now be written as:

$$q_t = \sum_{i=1}^{T} \hat{\delta}_{t,i}^{E} + \hat{\epsilon}_t,$$

where $\hat{\delta}_{t,i}^{E}$ represents the best update to the forecast of q_t that can be generated using the information available to the decision maker. The superscript "E" is used to denote echelon-based settings. In the SMI setting, when $\hat{\delta}_{t,i}^{E}$ is solely based on the signal $\delta_{t,i}^{s}$, we have $\hat{\delta}_{t,i}^{E} = (\theta_s + \theta_r\rho\sigma_{r,i}/\sigma_{s,i})\delta_{t,i}^{s}$. In the CFAR setting, $\hat{\delta}_{t,i}^{E} = \theta_r\delta_{t,i}^{r} + \theta_s\delta_{t,i}^{s}$ because both market signals $\delta_{t,i}^{r}$ and $\delta_{t,i}^{s}$ are known to the decision maker. We can therefore describe the demand evolution as one that depends on the most recent demand realization, and on the accumulated knowledge about future changes in demand, which is captured in the $\hat{\delta}^E$ signals. As before, it is sufficient to keep track of the aggregate value of the observed $\hat{\delta}^E$ signals that pertain to a specific random error q_t. We denote this sufficient statistic by $\hat{\Delta}_{t,i}^{E} \doteq \Sigma_{l=i}^{T} \hat{\delta}_{t,l}^{E}$. We can now define the $(T + 1)$-dimensional state vector X_t as follows:

$$\begin{aligned} X_t &\doteq (d_t - \mu, \hat{\Delta}_{t+1,1}^{E}, \hat{\Delta}_{t+2,2}^{E}, \ldots, \hat{\Delta}_{t+T,T}^{E})' \\ &= F_* X_{t-1} + W_t, \end{aligned} \tag{13}$$

and note that $d_t = G_* X_t + \mu$, with F^* and G^* defined as in Proposition 1. Each vector W_t is equal to $(\hat{\epsilon}_t, \hat{\delta}_{t+1,1}^{E}, \ldots, \hat{\delta}_{t+T,T}^{E})$, and it is useful to note that the vectors $\{W_t\}$ are statistically independent across t (see the proof of the next proposition). It follows that the demand evolution is of the same structure presented in §4.1, and in fact it has a simpler form because the decision maker fully observes the vector X_t during period t ($\Psi_t = X_t$ for all t). This observation allows us to propose the following forecasting process for our echelon-based settings:

Proposition 3 (SMI, CFAR). *The estimates of the leadtime demands, $\hat{D}_t^{(L)}$ and $\hat{D}_t^{(L+\tau)}$*, are given by

$$\begin{aligned} \hat{D}_t^{(L)} &= (L+1)\mu + \left(\sum_{l=1}^{L+1} \alpha^l\right)\cdot(d_{t-1} - \mu) \\ &\quad + \sum_{j=1}^{\min(T,L+1)} \left(\sum_{l=0}^{L+1-j} \alpha^l\right)\hat{\Delta}_{t+j-1,j}^{E}, \end{aligned} \tag{14}$$

with $\hat{D}_t^{(L+\tau)}$ being the same, replacing L with $L + \tau$ in (14). Furthermore, for all $t \geq T$, the uncertainty measure Ω_D associated with the MMSE estimate $\hat{D}_t^{(L)}$ is given by:

$$\begin{aligned} \Omega_D^{(L)} &\doteq \sum_{l=0}^{L} \left[G_*\left(\sum_{j=0}^{L-l} F_*^j\right)\Sigma_W\left(\sum_{j=0}^{L-l} F_*^j\right)' G_*'\right] \\ &= \mathrm{Var}(\hat{\epsilon}_t)\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l}\alpha^j\right)^2 \\ &\quad + \sum_{i=1}^{\min(L,T)} \mathrm{Var}(\hat{\delta}_{t,i}^{E})\cdot\left\{\sum_{l=0}^{L-i}\left(\sum_{j=0}^{L-i-l}\alpha^j\right)^2\right\}. \end{aligned} \tag{15}$$

The interpretation of (14) is very similar to that of (9) (see the remark after Proposition 1). How well the forecasting process performs, can be gauged by the mean-square error $\Omega_D^{(L)}$, and as can be seen from (15), the difference in performance between the SMI and the CFAR settings is purely due to the values of $\mathrm{Var}(\hat{\epsilon}_t)$ and $\mathrm{Var}(\hat{\delta}_{t,i}^{E})$. The following proposition investigates the difference between the value of $\Omega_D^{(L)}$ under the SMI and the CFAR settings.

Proposition 4. *The random variables $\hat{\epsilon}_t, \hat{\delta}_{t,1}^{E}, \ldots, \hat{\delta}_{t,T}^{E}$, are independent both in the SMI and the CFAR settings. In the SMI setting, $Var(\hat{\delta}_{t,i}^{E}) = (\theta_s\sigma_{s,i} + \rho\theta_r\sigma_{r,i})^2$, and $Var(\hat{\epsilon}_t) = \sigma_0^2 + (1 - \rho^2)\theta_r^2\, \Sigma_{i=1}^{T}\, \sigma_{r,i}^2$. In the CFAR setting, $Var(\hat{\delta}_{t,i}^{E}) = \theta_r^2\sigma_{r,i}^2 + 2\rho\theta_r\theta_s\sigma_{r,i}\sigma_{s,i} + \theta_s^2\sigma_{s,i}^2$, and $Var(\hat{\epsilon}_t) = \sigma_0^2$. Let $\Omega_{D,SMI}^{(L)}$ and $\Omega_{D,CFAR}^{(L)}$ be the values of $\Omega_D^{(L)}$ under the SMI and CFAR settings, respectively. Then,*

$$\begin{aligned} &\Omega_{D,SMI}^{(L)} - \Omega_{D,CFAR}^{(L)} \\ &\quad = (1-\rho^2)\sum_{i=1}^{T}(\theta_r\sigma_{r,i})^2 \\ &\quad\quad \times \left\{\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l}\alpha^j\right)^2 - \mathbf{1}\{i \leq L\}\sum_{l=0}^{L-i}\left(\sum_{j=0}^{L-i-l}\alpha^j\right)^2\right\} \\ &\quad \geq 0. \end{aligned} \tag{16}$$

($\mathbf{1}\{\cdot\}$ is the indicator function.) As expected, the mean-square error is lower in the CFAR setting than in the

SMI setting because market signals that are observable by the retailer are not taken into account in SMI. This explains why the expression in (16) consists of the terms $(1 - \rho^2)(\theta_r \sigma_{r,i})^2$—these terms are exactly the variance of the conditional distribution $(\theta_r \delta^r_{t,i} + \theta_s \delta^s_{t,i}) \mid \delta^s_{t,i}$. In other words, they represent the portions of the variance that can be explained *only* by the retailer. Therefore, the difference in (16) is expected to grow with θ_r—i.e., when the retailer's observations are useful in explaining future demand, and/or when ρ is smaller—i.e., when the retailer's market signals provide information that cannot be statistically learned from the supplier's signals. Furthermore, it can be shown the difference increases with α; see further discussion in §6.2.

5. Policy Coordination

This section describes a simulation-based method to search for the base-stock levels that minimize the long-run total supply chain costs, for each of our three settings. Let y^m_t be the *net inventory*[6] of member m at the end of period t, and consider the N-period average systemwide cost $(1/N)\ \Sigma^N_{t=1} [c^s(y^s_t) + c^r(y^r_t)]$, where $c^m(y) \doteq h^m \cdot \max(y, 0) - p^m \cdot \min(0, y)$, and $p^s = 0$. Because we are interested in long-run performance (i.e., $N \to \infty$), we can focus our analysis on the following cost function:

$$C_N \doteq \frac{1}{N} \sum_{t=1}^{N} [c^s(y^s_t) + c^r(y^r_{t+L})].$$

Depending on whether we deal with installation-based or echelon-based settings, our purpose is to minimize the term $\lim_{N\to\infty} C_N$, by an appropriate choice of the parameters (γ^r, γ^s) or (Γ^r, Γ^s), respectively. We start with an analysis of the LMI setting (§5.1), and then continue with the echelon-based inventory systems (§5.2).

5.1. Policy Coordination in the LMI Setting

Given our assumption that $\text{Std}(q_t)/\mu$ is small, we can comfortably assume that the retailer's orders satisfy[7] $A_t = \beta^r_t - \beta^r_{t-1} + d_{t-1}$, and that the supplier's inventory position at the beginning of every period t is equal to β^s_t. It hence follows that the supplier's *net* inventory at the end of period t is equal to $y^s_t = \beta^s_{t-\tau} - \Sigma^\tau_{l=1} A_{t-\tau+l} = \gamma^s - (\Sigma^\tau_{l=1} A_{t-\tau+l} - \hat{A}^{(\tau)}_{t-\tau})$. Next, observe that the retailer's net inventory at the end of period $t + L$ is given by the retailer's inventory position at the beginning of period t (after he has placed his order), minus all deliveries to the retailer that are backlogged due to supplier's shortages, minus the total demand during periods $t, t + 1, \ldots, t + L$. In other words, $y^r_{t+L} = \beta^r_t - \max(R^s_{t-\tau} - \gamma^s, 0) - \Sigma^L_{l=0} d_{t+l} = \gamma^r - \max(R^s_{t-\tau} - \gamma^s, 0) - (\Sigma^L_{l=0} d_{t+l} - \hat{D}^{(L)}_t)$. We define $R^s_{t-\tau} \doteq \Sigma^\tau_{l=1} A_{t-\tau+l} - \hat{A}^{(\tau)}_{t-\tau}$ as the supplier's forecasting error for period $t - \tau$ and $R^r_t \doteq \Sigma^L_{l=0} d_{t+l} - \hat{D}^{(L)}_t$ as the retailer's forecasting error for period t. Both of these forecasting error terms are independent of the policy parameters γ^r and γ^s. Our problem can be presented as follows:

$$\min_{\gamma^s} \Big\{ \mathsf{E}[c^s(\gamma^s - R^s)] + \min_{\gamma^r} \mathsf{E}[c^r(\gamma^r - \max(R^s - \gamma^s, 0) - R^r)] \Big\} \quad (17)$$

where (R^r, R^s) is a pair of binormally distributed random variables that represent the joint long-run distribution of the process $\{(R^r_t, R^s_{t-\tau}) : t \geq 1\}$. The joint distribution of this pair is given in the following proposition.

PROPOSITION 5. *In the LMI setting, the joint distribution of R^r, R^s is binormal, with mean zero. Specifically, $Var(R^r) = \Omega^{LMI}_D$, $Var(R^s) = \Omega^{LMI}_A$; see §§4.2.1 and 4.2.2, respectively, and*

$$\text{Cov}(R^r, R^s) = P\tilde{F}^{\tau-1}\Omega^Z \Big(\sum_{l=0}^{\tau-1} \tilde{F}^l\Big)' \tilde{G}' + \sum_{l=0}^{\tau-2} \Big\{ P\tilde{F}^l \Sigma_{\tilde{V}} \Big(\sum_{j=0}^{l} \tilde{F}^j\Big)' \tilde{G}' \Big\}$$

with all matrices in the last equation being adopted from §4, and $P \doteq G(\Sigma^L_{l=0} F^l)(F - FKH) \cdot (\mathbf{0}_{(2T+1)\times(T+1)}, I_{(2T+1)\times(2T+1)})$.

We propose the following simulation-based algorithm to find the optimal values of γ^r and γ^s: A large set of random values $\{(u^r_n, u^s_n), n = 1, \ldots, N\}$ is generated from the joint distribution of (R^r, R^s). The in-

[6]Net inventory is the on-hand inventory minus backlogs.
[7]See, e.g., Aviv (2001b) and Lee et al. (2000).

side minimization term in (17) is essentially a newsvendor problem with γ^r being the ''order quantity'' and $R^r + \max(R^s - \gamma^s, 0)$ being the unknown ''demand.'' Hence, for any given value of γ^s, we pick the minimal value of γ^r such that $-p^r + (h^r + p^r)(1/N) \sum_{n=1}^{N} \mathbf{1}\{u_n^r + \max(u_n^s - \gamma^s, 0) \cdot \gamma^r\} \geq 0$. We use the bisection method for this purpose. The minimization of the entire expression in (17) is done by a line search on γ^s. Finally, the expected cost under the best choice of policy parameters γ^{*r} and γ^{*s} is given by

$$\frac{1}{N}\sum_{n=1}^{N}[c^s(\gamma^{*s} - u_n^s) + c^r(\gamma^{*r} - \max(u_n^s - \gamma^{*s}, 0) - u_n^r)].$$

This final cost estimation procedure can be repeated enough times, with different streams of $\{(u_n^r, u_n^s), n = 1, \ldots, N\}$, so as to generate a desirable statistical confidence interval.

5.2. Echelon-Based Settings

Let y_t^m be defined as before, with backlogs at the supplier's level being the amount by which the retailer's echelon inventory position is lower than the target level B_t^r. The supplier's net inventory at the beginning of period t, after deliveries to the retailer are made for this period, is equal to:[8]

$$\begin{aligned} y_t^s &= B_{t-\tau}^s - B_{t-\tau}^r - \sum_{l=1}^{\tau}(B_{t-\tau+l}^r - B_{t-\tau-1+l}^r + d_{t-\tau-1+l}) \\ &= (\Gamma^s - \Gamma^r) - \left(\sum_{l=0}^{L+\tau} d_{t-\tau+l} - D_{t-\tau}^{(L+\tau)}\right) + \left(\sum_{l=0}^{L} d_{t+l} - D_t^{(L)}\right) \\ &= (\Gamma^s - \Gamma^r) - R_{t-\tau}^s + R_t^r. \end{aligned}$$

The component $(B_{t-\tau+l}^r - B_{t-\tau-1+l}^r + d_{t-\tau-1+l})$ in the last equation is basically the ''new'' desired delivery quantity (i.e., excluding previous backlogs) in period $t - \tau + l$. The terms $R_t^r \doteq \sum_{l=0}^{L} d_{t+l} - D_t^{(l)}$ and $R_{t-\tau}^s \doteq \sum_{l=0}^{L+\tau} d_{t-\tau+l} - D_{t-\tau}^{(L+\tau)}$ represent two forecasting errors, each corresponding to a different forecasting time and a different lead time. Next, the lagged retailer's net inventory term y_{t+L}^r is simply given by the actual inventory position of the retailer in period t, minus the total demand during periods $t, t + 1, \ldots, t + L$. In other words,

$$\begin{aligned} y_{t+L}^r &= B_t^r + \min(y_t^s, 0) - \sum_{l=0}^{L} d_{t+l} \\ &= \Gamma^r - R_t^r + \min(\Gamma^s - \Gamma^r - R_{t-\tau}^s + R_t^r, 0). \end{aligned}$$

Our problem can be presented in the following form:

$$\min_{\delta}\Big\{\mathsf{E}[c^s(\delta - \tilde{R}^s)] + \min_{\Gamma^r}\mathsf{E}[c^r(\Gamma^r - \max(\tilde{R}^s - \delta, 0) - R^r)]\Big\} \quad (18)$$

where $\delta \doteq \Gamma^s - \Gamma^r$, and $(R^r, \tilde{R}^s)$ is a generic pair of random variables that represents the joint distribution of $(R_t^r, R_{t-\tau}^s - R_t^r)$ for every t. As can be seen, the structure of problem (18) is equivalent to the structure of the installation-based problem (17). Hence, we propose the same algorithm and cost estimation procedure as §5.1 to find the optimal values of Γ^{*r} and δ^*. To this end, we need to characterize the joint distribution of $(R^r, \tilde{R}^s)$:

Proposition 6. *Under both SMI and CFAR settings, the random variables R^r and $\tilde{R}^s$ are independent and normally distributed. Their means are equal to zero, and their variances are $Var(R^r) = \Omega_D^{(L)}$ and $Var(\tilde{R}^s) = \Omega_D^{(L+\tau)} - \Omega_D^{(L)}$. The values of $\Omega_D^{(\cdot)}$ are determined by (15), with the value of Σ_W being set according to the specific setting; see Proposition 4.*

We finally remark that because the correlation between R^r and $\tilde{R}^s$ is zero, problem (18) is equivalent to the Clark and Scarf (1960) problem in a stationary, i.i.d. demand environment. Therefore, the value of Γ^{*r} is given by the newsvendor value

$$\Gamma^{*r} = \Phi^{-1}\left(\frac{h^s + p^r}{h^r + p^r}\right) \cdot \sqrt{\Omega_D^{(L)}}.$$

(The function Φ^{-1} is the inverse of the standard normal distribution function.) This, however, helps us achieve only a small reduction in the computation time. We also refer the reader to §4 of Cachon and Zipkin (1999) for a concise description of a similar optimization algorithm for this case.

[8]We again use the assumption that $\mathrm{Std}(q_t)/\mu$ is small.

6. Computational Study and Managerial Insights

We have conducted a numerical study to analyze our three settings over a large variety of system parameters. For each combination of parameters, we have used our models to assess the best average cost for the supply chain, under LMI, SMI, and CFAR. For each setting investigated, we simulated 10,000 random values $\{(u_n^r, u_n^s)\}$, from the joint distribution of (R^r, R^s) (or $R^r, \tilde{R}^s$ in the case of the echelon-based settings) and solved problem (17) (or problem (18)). We then estimated the expected costs by repeating the simulation enough times so as to generate a 95% confidence interval with a range not exceeding 1.5% of the estimated average cost. Throughout our study, we assume $T = 4$, $\text{Var}(q_t) = 10$, and $\sigma_{r,i}^2 = \sigma_{s,i}^2 = \sigma_i^2$ for every i. We also keep the pattern $\{\sigma_i\}$ the same across all instances, up to constant multiplication: Specifically, we assume that $\sigma_i^2 = 0.8\sigma_1^2, 0.6\sigma_1^2, 0.4\sigma_1^2$, for $i = 2, 3, 4$, and $\sigma_i^2 = 0$ for all $i > 4$. In addition, we assume $h^r = 1$, $p^r = 19$, and $h^s = 0.5$. We have experimented with different values of p^r, but the results do not add much to our discussion below, and hence they are not reported.

6.1. The Strength of the Explanatory Market Signals

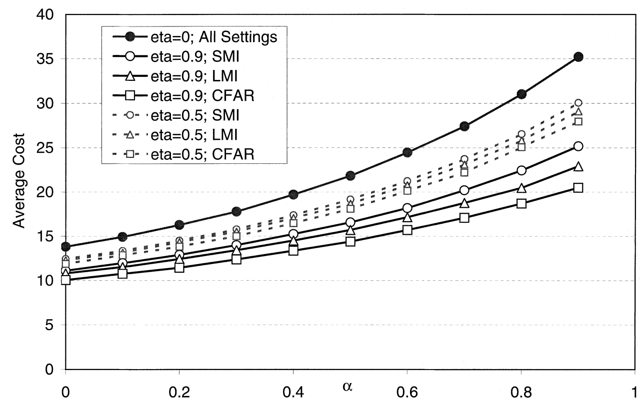
We begin our numerical study with an investigation of the combined impact of two main characteristics of the demand model on system performance: First, we consider α, the degree of auto-correlation between demands in consecutive periods. Second, we consider the *strength of the explanatory market signals*, or the portion of the variance of q_t that can be explained by the supply chain members using *all* market signals observed by *both* of them up to time t. This portion is hence defined as

$$\eta \doteq 1 - \text{Var}(\epsilon_t)/\text{Var}(q_t)$$

$$= \sigma_1^2 \cdot (\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2) \cdot \frac{\sum_{i=1}^{T} (\sigma_i^2/\sigma_1^2)}{\text{Var}(q_t)}.$$

(See Equation (3).) In general, the value of η provides only a crude way for describing the potential ability to learn about the variables q_t from early market signals, because it does not indicate whether this portion of variability is explainable progressively over time, or perhaps only at the last moment. However, the use of this parameter is appropriate in our study because we keep the *pattern* of $\{\sigma_i\}$ the same across all instances. In particular, we vary the value of η in the range [0, 1], through an appropriate choice of σ_1^2. Specifically, in view of the expression above, this is done through $\sigma_1^2 = \eta\ \text{Var}(q_t)/[(\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2) \sum_{i=1}^{T} (\sigma_i^2/\sigma_1^2)]$. (Observe that $\sum_{i=1}^{T} (\sigma_i^2/\sigma_1^2)$ is constant.) For instance, $\eta = 0$ represents the case in which no market signals are available to the decision makers, and the case $\eta = 1$ represents the case in which market signals exist and allow the members to perfectly predict the value of q_t at the beginning of period t, given that they share the values of the market signals they have observed. We consider the following combination of parameters: $L, \tau \in \{1, 2, 3\}$, $\rho = 0.5$, $\theta_r = \theta_s = \theta$, and $\alpha \in \{0, 0.1, \ldots, 0.9\}$. We then varied the value of $\eta \in \{0, 0.25, 0.5, 0.75, 0.9\}$, by setting σ_1^2 according to the expression above, and $\sigma_0^2 = \text{Var}(q_t) \cdot (1 - \eta)$. Overall, 450 instances were examined in this initial study.

Figure 1 **Impact of α on Average Supply Chain Cost Performance ($L = \tau = 2$, $\eta = 0, 0.5, 0.9$)**



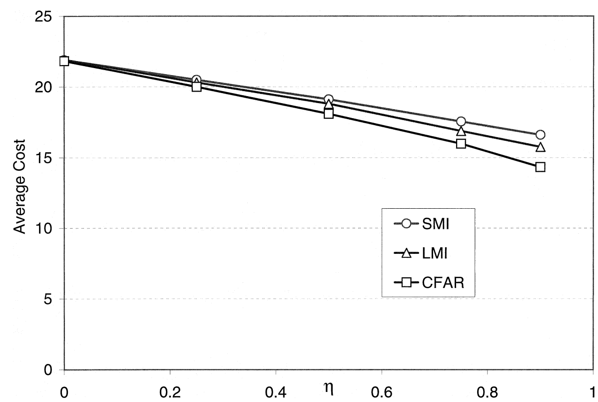
Two special cases of our model are the AR(1) model examined by Lee et al. (2000), which is equivalent to the case $\eta = 0$, and the model discussed by Aviv (2001a), which is a variant of the case $\alpha = 0$. Figure 1 shows the average supply chain cost as a function of α for each of our three settings, for the cases $L = \tau = 2$, and $\eta = 0, 0.5, 0.9$ only. However, all of the statements we make below with regard to this figure are supported by the rest of the experimental set as well. We observe that in the case $\eta = 0$, all settings perform *identically.* This result is clear when comparing SMI with CFAR, because both of these settings are centrally managed with the same amount of information available to the decision makers. But, the equivalence of LMI and CFAR (or SMI) is less straightforward.

PROPOSITION 7. *When $\theta_r = \theta_s = 0$ or, alternatively, when $\sigma_i = 0$ for all i, the supply chain gains no benefits by providing the supplier with direct access to demand and inventory information at the retailer's level.*

(See Aviv 2001b or Raghunathan 2001 for a proof.) The proposition stands in contrast to Lee et al. (2000), but is supported by Aviv (2001b) and Raghunathan (2001). Raghunathan explains that Lee et al. observe benefits of information sharing in their model because the forecasts they developed for the supplier do not use all of the historic order data available to the supplier. When the supplier uses all of the information available to him in the construction of MMSE estimates, the value of information sharing drops to zero. As a consequence of this theoretical result, we augment Lee et al.'s comment, by arguing that even the AR(1) model may be insufficient to explain the benefits of information sharing practices (e.g., in the grocery industry) because it does not capture the existence of advanced demand information. Our paper deals precisely with this type of model extension. In fact, Raghunathan suggested in the conclusion of his paper that: "Information sharing will be useful in Lee et al.'s model under the following demand model, $D_t = d + \alpha D_{t-1} + \gamma X_{t-1} + \epsilon_t$, in which X_t is a retailer action ... taken during period t that affect the demand during the next period." Raghunathan's suggestion is indeed supported by our results for the cases in which $\eta > 0$. For instance, Figure 1 shows that when $\eta = 0.5, 0.9$, performance gaps exist between the three settings, even in the case $\alpha = 0$.

Continuing with the same experimental set described above, Figure 2 depicts the average costs as functions of η for $\alpha = 0.5$ and $L = \tau = 2$. First, and most clearly, the costs are declining in the value of η, or as the explainable portion of the error terms (q_t) increases. Second, this figure shows that the magnitude of the performance gaps between our settings becomes larger when more advanced demand information prevails in the system. The following proposition generalizes the latter observation to all possible combinations of the parameters $\{\theta_r, \theta_s, \rho\}$, with regard to the cost difference between the SMI and CFAR settings.

Figure 2 **Impact of the "Explainable" Portion η on Average Supply Chain Cost Performance** ($L = \tau = 2$, $\alpha = 0.5$)



PROPOSITION 8. *Assume that the pattern $\{\sigma_i\}$ is fixed up to constant multiplication, and let V_S^r and V_S^s be the values of $Var(R^r)$ and $Var(\tilde{R}^s)$, respectively, in the SMI setting. Similarly, define V_C^r and V_C^s for the CFAR setting. Then, the following terms are increasing in η for all possible values of θ_r, θ_s, and ρ: $\sqrt{V_S^r} - \sqrt{V_C^r}$, and $\sqrt{V_S^r + V_S^s} - \sqrt{V_C^r + V_C^s}$.*

Consider the difference terms in Proposition 8 as proxies for the performance gaps between the SMI and CFAR settings: $\sqrt{V_S^r} - \sqrt{V_C^r}$ for costs at the retailer level and $\sqrt{V_S^r + V_S^s} - \sqrt{V_C^r + V_C^s}$ for the total supply chain costs. This proposition suggests that as companies improve their ability to explain future demands on the basis of early market signals, the potential of CFAR vis-à-vis SMI increases. Hence, a standard VMI program (i.e., no collaborative forecasting) that has worked well up until a certain point of time may need to be replaced by CFAR as the trading partners improve their forecasting performance. As a conjecture, we expect that as technology advances in industry (e.g., making data mining solutions more affordable and powerful), we shall see CFAR replacing most standard VMI programs. An exception to that would be in settings in which the supplier's forecasting capabilities are substantially better than those of the retailer.

6.2. Early Market Signals and the Parameter α

In agreement with earlier results by Lee et al. (2000) and Graves (1999),[9] we see that the costs increase considerably with α, i.e., as demands in consecutive periods are more correlated. For the case $\alpha = 0$, Aviv (2001a) has illustrated that significant benefits can be achieved by taking into account early demand signals. It seems that when $\alpha > 0$, the sizes of these benefits are even larger. For example, in the instances of Figure 1, if $\eta = 0.9$, but the supply chain members do not use their explanatory variables, then their cost would be given by the curve of $\eta = 0$. When they incorporate the explanatory variables into their (say) LMI setting, they would be able to gain percentage benefits of (22%, 25%, 30%, 35%), for $\alpha = (0, 0.3, 0.6, 0.9)$, respectively.

But why should a company be concerned about the value of taking into account early demand signals? The answer to this has to do with the *ability* of the company to make use of market signals. Consider, for example, Longs Drug Stores, a major United States drug chain with \$3.7 billion in annual sales. Lee and Whang (2001) report that although the pharmacists and buyers of this chain were intimately familiar with seasonal patterns of demand for major drugs, they lacked the data and decision-support systems that would enable them to effectively manage the enormous number of stock-keeping units sold by the company. It is therefore to the advantage of such companies to be able to use models for gauging the potential benefits that decision-support technology can provide them with. Longs Drug Stores partnered with a third-party, demand-management solution provider, Nonstop Solutions, that offers state-of-the-art methodologies to optimize their demand management process. Among the strengths provided by Nonstop Solutions is the capability to analyze rich and timely data to generate forecasts.

[9]Graves investigated a different type of time-series demand model, in which $d_t = d_{t-1} - (1 - \alpha)\epsilon_{t-1} + \epsilon_t$.

6.3. Relative Explanation Power

Intuitively, among the three settings, CFAR should have the best inventory-cost performance. Furthermore, when one compares the SMI and LMI settings, it is common sense that when the supplier cannot explain much of the demand uncertainty, the SMI setting would be less beneficial, and perhaps even result in a loss[10] when it replaces LMI, as many companies have experienced in their implementation of VMI programs. For example, the supplier's lack of forecasting ability was one of the reasons that a large consumer electronics company was dissatisfied with several VMI programs it implemented in 1995 (see other examples in Mathews 1995 and Fiddis 1997). We therefore argue that the relative ability of the supply chain members to use early market signals for explaining future demand is a determining factor in the cost benefit of SMI. In fact, our cost analysis above demonstrates that systemwide costs are purely determined by the characteristics of the forecasting errors associated with the estimates of lead-time demands, as well as the interaction between them; see (17) and (18).

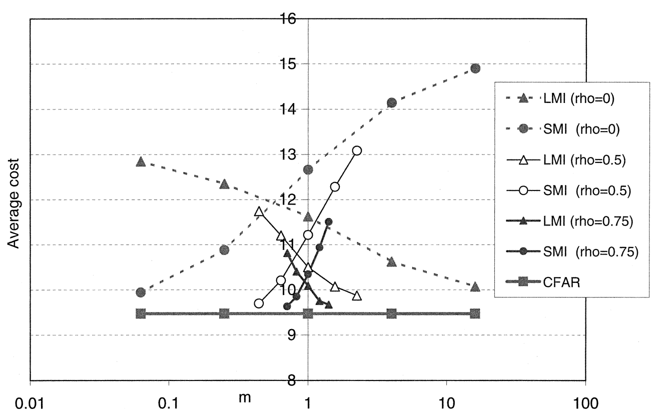
Recall, that the portions of $\mathrm{Var}(q_t)$ that can be explained through $\delta^r_{t,i}$ and $\delta^s_{t,i}$ are $(\theta_r + \rho\theta_s)^2\sigma_i^2$ and $(\rho\theta_r + \theta_s)^2\sigma_i^2$, respectively. We hence define the measure

$$m \doteq \left(\frac{\theta_r + \rho\theta_s}{\rho\theta_r + \theta_s}\right)^2 \tag{19}$$

as the *relative explanation power of the retailer*—in other words, the extent to which the market signals observed directly by the retailer are more informative than those observed directly by the supplier. As can be seen from (19), the value of m depends on θ_s and θ_r through the ratio $w \doteq \theta_s/\theta_r$; i.e., $m = [(1 + \rho w)/(\rho + w)]^2$. We have studied an additional set of 150 instances as follows: $L = \tau = 1$, $\sigma_1^2 = 1$, $\alpha \in \{0, 0.1, \ldots, 0.9\}$, $\rho \in \{0, 0.5, 0.75\}$, and $w \in \{0.25, 0.5, 1, 2, 4\}$. For every value of w, we set the value of θ_r (and accordingly $\theta_s = w\theta_r$) so that $\sigma_0^2 = 2.5$. Keeping the value of σ_0^2 fixed means that the portion of the variance of q_t that can be explained through *both* streams of signals, δ^r and δ^s, is constant across m. Hence, any change in performance across different values of m

[10]Figure 1 serves as an *example* in which SMI under performs LMI.

Figure 3 **Impact of the Retailer's Relative Explanation Power (m) on Average Supply Chain Cost Performance ($L = \tau = 1$, Var(q_t) = 10, $\sigma_0^2 = 2.5$, $\alpha = 0.5$)**



would be purely due to the *relative* explanation power of the supply chain members. Figure 3 below shows, for the case $\alpha = 0.5$, the average cost of the supply chain as a function of m for each one of our three settings (the horizontal axis is presented in a logarithmic scale).

We first note that because the combined explanatory power of the retailer's and the supplier's observed market signals is constant across m, the cost under CFAR is also independent of m. This is represented by the straight line in Figure 3. In addition, Figure 3 depicts three pairs of curves, corresponding with different values of $\rho \in \{0, 0.5, 0.75\}$.[11] Each pair of curves shows the average supply chain cost under SMI and LMI, respectively. As expected, we observe that the cost under SMI increases sharply as the supplier's relative explanation power declines, and that there exists a threshold for m, below which LMI performs better than SMI. We shall denote this correlation-dependent threshold value by $m^*(\rho)$.

The cost under LMI seems to depend to a larger extent on the retailer's than on the supplier's explanation power. One possible reason for this is that forecasting errors are more critical at the retailer's level. In our model, this is reflected by the high direct penalty charges for shortages at the retailer level. The rationale behind this is that a shortage at the retailer level is usually more harmful to the supply chain (e.g., due to loss of customer satisfaction) than a delay of delivery of a unit from the supplier to the retailer, which only occasionally may cause a shortage later at the retailer site. A second reason is that under LMI, the supplier is still able to extract information about future order quantities from the history of the retailer's orders. Along the same lines, we observe in Figure 3 (as well as in the rest of our results), that the point at which the LMI and SMI cost curves cross is less than 1; i.e., $m^*(\rho) \leq 1$. This may be supported by the same intuition described above, and the practical meaning of this observation is important—it suggests that in an implementation of VMI programs, it is crucial to ascertain that the supplier will be capable of observing and incorporating early demand signals that are at least as informative as those observed by the retailer;[12] otherwise, collaborative forecasting may be necessary, and if not justified, LMI may be the best choice for the supply chain.

Another important point that Figure 3 raises is concerned with the correlation parameter $\rho = \mathrm{Corr}(\delta^r_{t,i}, \delta^s_{t,i})$. We observe (again, through more experimentation) that when ρ increases, the cost at the crossing point of the SMI and LMI curves decreases. In fact, as ρ grows closer to one, the curves' crossing point $m^*(\rho)$ approaches one, and the cost of this meeting point approaches that of the CFAR setting. This asymptotic result is clear because the case $\rho = 1$ means that the retailer and the supplier virtually observe the *same* explanatory variable terms $\theta_r\delta^r_{t,i} + \theta_s\delta^s_{t,i}$. Hence, CFAR becomes equivalent to SMI and LMI.

The observations made above seem to imply that when the supplier's relative explanation power is large (small m), the SMI setting may be the preferable choice for the supply chain. (Note that the costs under CFAR and SMI are the same when m is very small.) When the relative explanation power of the retailer and the supplier are similar (around $m = 1$ and below), the answer *depends on the value of* ρ: If ρ is small, CFAR may be the best choice, whereas if ρ is high,

[11]Recall that when $\rho = 1$, the SMI and LMI settings are equivalent to CFAR.

[12]This proposition is made in the context of the system discussed in this paper; i.e., under no capacity constraints or economies of scale.

the three settings will have similar costs. Finally, when the retailer's relative explanation power is large, the choice should be between LMI and CFAR, and it depends on the gap between these two curves. Here—in contrast to the case of small values of m (where the SMI and CFAR setting are alike)—the relative performance of LMI, compared to CFAR, depends on how well the supplier can learn from the history of retailer's orders. Figure 3 seems to suggest that the gap between LMI and CFAR drops rapidly as ρ increases.

7. Conclusion

We have developed a stylized framework to describe a two-level supply chain that faces an auto-regressive demand process. We extend the current literature (particularly, the closely related paper of Lee et al. 2000) by modeling the ability of companies to observe early market signals, and hence improve their forecasting performance. This extension seems to be valuable because it allows us to study the benefits of information sharing in the realistic, non-i.i.d. demand environment. By doing so, this work also adds a sharper focus to the study conducted by Aviv (2001a). We have demonstrated through numerical examples that the consideration of VMI and CFAR programs becomes more important as the demand process is more correlated across periods, and as companies are able to explain a larger portion of the demand uncertainty through early demand information. In addition, we have argued that the choice of the best system depends greatly on the understanding of the interaction between the explanatory power of the supply chain members.

Acknowledgments

The author greatly acknowledges the comprehensive and thorough feedback received from an anonymous referee, the Senior Editor, and the Editor-in-Chief. This paper has considerably benefited from their valuable input.

Appendix. Proofs

Proof of Proposition 1. Define the vector $Y_t' = (d_t - \mu, \hat{\Delta}^r_{t+1,1}, \hat{\Delta}^r_{t+2,2}, \ldots, \hat{\Delta}^r_{t+T,T}) \in \mathbf{R}^{T+1}$. Clearly, this state vector is fully observable by the retailer *during* period t (albeit, after the replenishment decision for this period is made). Furthermore, Y_t follows the evolution $Y_t = F^*Y_{t-1} + V^*_t$, where

$$V_{*t} = \left(\epsilon_t + \theta_s\left(\Delta^s_{t,1} - \rho \sum_{i=1}^{T} \frac{\sigma_{s,i}}{\sigma_{r,i}} \delta^r_{t,1}\right)\hat{\delta}^r_{t+1,1} \cdots \hat{\delta}^r_{t+T,T}\right)'.$$

Next, it can be easily verified that the vectors $\{V_t\}$ are i.i.d. and that the covariance matrix of V_t is Σ_*. Finally, note that

$$G_* \sum_{l=1}^{L+1} F^l_* = \begin{cases} \left(\sum_{l=1}^{L+1} \alpha^l, \sum_{l=0}^{L} \alpha^l, \sum_{l=0}^{L-1} \alpha^l, \ldots, 1, \underbrace{0, \ldots, 0}_{(T-L-1)\,\text{entries}}\right) & \text{if } L < T, \\ \left(\sum_{l=1}^{L+1} \alpha^l, \sum_{l=0}^{L} \alpha^l, \sum_{l=0}^{L-1} \alpha^l, \ldots, \sum_{l=0}^{L+1-T} \alpha^l\right) & \text{if } L \geq T. \end{cases}$$

Equation (9) now follows. Because the state vector Y_t is fully observable by the decision maker (albeit, after the beginning of period t), we have $\hat{Y}_{t\,|\,t-1} = F^*Y_{t-1}$. The value of Ω^{LMI}_D follows now directly from Aviv (2001b).

Proof of Proposition 2. The dynamics equation $Z_t = \tilde{F}Z_{t-1} + \tilde{V}$ and the last equation of the proposition are straightforward. Because $\mathrm{E}[\Delta^s_{t,i} \mid \Delta^r_{t,i}] = \rho\Delta^r_{t,i}$, it is easily verified that

$$\begin{aligned} \hat{X}_{t|t-1} = (&\alpha(d_{t-1} - \mu) + (\theta_r + \rho\theta_s)\Delta^r_{t,1}, \Delta^r_{t+1,2}, \ldots, \Delta^r_{t+T-1,T}, \\ &0, \rho\Delta^r_{t+1,2}, \ldots, \rho\Delta^r_{t+T-1,T}, 0). \end{aligned}$$

Therefore, $\hat{X}_{t\,|\,t-1} = FKHX_{t-1}$; see the description of FK in the proof of Proposition 1. Next, note by (6) that $A_{t+1} = G(\Sigma_{l=0}^{L} F^l)FKHX_t - G(\Sigma_{l=0}^{L} F^l)FK(HX_{t-1}) + GX_t + \mu$. The value of the $(2T + 1)$-dimensional vector $G(\Sigma_{l=0}^{L} F^l)$ is given by

$$\left[G\left(\sum_{l=0}^{L} F^l\right)\right]_i = \begin{cases} \sum_{l=0}^{L} \alpha^l & i = 0, \\ \theta_r \sum_{l=0}^{L-i} \alpha^l & 1 \leq i \leq \min(L, T), \\ \theta_s \sum_{l=0}^{L+T-i} \alpha^l & T + 1 \leq i \leq T + \min(L, T), \\ 0 & \text{otherwise.} \end{cases}$$

(The index i ranges from 0 to $2T$.) A multiplication of $G(\Sigma_{l=0}^{L} F^l)$ by FK yields the $(T + 1)$-dimensional vector M. This justifies the equation $\Psi^s_t = \tilde{H}Z_t$.

Proof of Proposition 3. The proof is similar to that of Proposition 1. To verify (15), observe that:

$$G_*\left(\sum_{j=0}^{x} F^j_*\right) = \begin{cases} \left(\sum_{j=0}^{x} \alpha^j, \sum_{j=0}^{x-1} \alpha^j, \ldots, 1, 0, \ldots, 0\right) & \text{if } x < T, \\ \left(\sum_{j=0}^{x} \alpha^j, \sum_{j=0}^{x-1} \alpha^j, \ldots, \sum_{j=0}^{x-T} \alpha^j\right) & \text{if } x \geq T. \end{cases}$$

The rest follows by simple matrix algebra; recall the Σ_W is diagonal.

Proof of Proposition 4. The proof of the first part is straightforward for the CFAR case because of the independence of $\delta^m_{t,i}$ across t and i and because the latter are independent of ϵ_t. The variance of each of the $\hat{\delta}$ variables can be obtained by simple algebra. In the

SMI case, the independence between the $\hat{\delta}^E$ variables is trivial. To show that $\hat{\epsilon}_t$ is independent of $\hat{\delta}^s_{t,i}$ for all $i = 1, \ldots, T$, observe that

$$\begin{aligned}\mathrm{Cov}(\hat{\delta}^s_{t,i}, \hat{\epsilon}_t) &= (\theta_s + \theta_r\rho\sigma_{r,i}/\sigma_{s,i})\theta_r \\ &\quad \div \sigma_{s,i}[\sigma_{s,i}\mathrm{Cov}(\delta^s_{t,i}, \Delta^r_{t,1}) - \rho\sigma_{r,i}\mathrm{Cov}(\delta^s_{t,i}, \Delta^s_{t,1})] \\ &= (\theta_s + \theta_r\rho\sigma_{r,i}/\sigma_{s,i})\theta_r \\ &\quad \div \sigma_{s,i}[\sigma_{s,i}\mathrm{Cov}(\delta^s_{t,i}, \delta^r_{t,i}) - \rho\sigma_{r,i}\mathrm{Var}(\delta^s_{t,i}, \delta^s_{t,i})] = 0.\end{aligned}$$

To verify (16), simply note that the difference in $\mathrm{Var}(\epsilon_t)$ between the SMI and CFAR settings is $(1 - \rho^2)\theta_r^2 \sum_{i=1}^T \sigma_{r,i}^2$ and that the difference for each $\mathrm{Var}(\hat{\delta}^E_{t,i})$-value is $(1 - \rho^2)\theta_r^2\sigma_{r,i}^2$. The proof then follows directly from (15).

Proof of Proposition 5. Recall that in the LMI case, $\hat{X}_{t\,|\,t-1} = FKHX_{t-1}$, and hence by Proposition 2: $\hat{X}_{t\,|\,t-1} = FKH\cdot(\mathbf{O}_{(2T+1)\times(T+1)}, I_{(2T+1)\times(2T+1)})\cdot Z_{t-1}$. Therefore, because $X_t = FX_{t-1} + V_t = F(\mathbf{0}_{(2T+1)\times(T+1)}, I_{(2T+1)\times(2T+1)})Z_{t-1} + V_t$,

$$\begin{aligned}R^r_t &= G\left(\sum_{l=0}^{L} F^l\right)(X_t - \hat{X}_{t|t-1}) + G\sum_{l=1}^{L}\left(\sum_{j=0}^{L-l} F^j\right)V_{t+l} \\ &= PZ_{t-1} + G\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l} F^j\right)V_{t+l}.\end{aligned}$$

By expanding the term Z_{t-1} we obtain

$$R^r_t = P\tilde{F}^{\tau-1}Z_{t-\tau} + \sum_{l=1}^{\tau-1} P\tilde{F}^{\tau-1-l}\tilde{V}_{t-\tau+l} + G\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l} F^j\right)V_{t+l}.$$

Next, note that

$$R^s_{t-\tau} = \tilde{G}\left(\sum_{l=0}^{\tau-1}\tilde{F}^l\right)(Z_{t-\tau} - \hat{Z}_{t-\tau|t-\tau+1}) + \sum_{l=1}^{\tau-1}\tilde{G}\left(\sum_{j=0}^{\tau-1-l}\tilde{F}^j\right)\tilde{V}_{t-\tau+l}.$$

The third component in the expression for R^r_t is independent of R^s_t, and the first (second) component of R^r_t is not correlated with the second (first) component of $R^s_{t-\tau}$. Thus,

$$\mathrm{Cov}(R^r_t, R^s_{t-\tau}) = P\tilde{F}^{\tau-1}\Omega^Z\left(\sum_{l=0}^{\tau-1}\tilde{F}^l\right)'\tilde{G}' + \sum_{l=1}^{\tau-1}\left\{P\tilde{F}^{\tau-1-l}\Sigma_{\tilde{V}}\left(\sum_{j=0}^{\tau-1-l}\tilde{F}^j\right)'\tilde{G}'\right\}.$$

Proof of Proposition 6. Because the state vector X_{t-1} is fully observable at point t, we have $X_t - \hat{X}_{t\,|\,t-1} = W_t$. Thus, $R^r_t = G_* \sum_{l=0}^{L}(\sum_{j=0}^{L-l} F^j_*)W_{t+l}$, and $R^s_{t-\tau} = G_* \sum_{l=0}^{L+\tau}(\sum_{j=0}^{L+\tau-l} F^j_*)W_{t-\tau+l}$. But the W_t components are independent across time, and hence $\mathrm{Var}(R^r_t) = \Omega_D^{(L)}$, $\mathrm{Var}(R^s_{t-\tau}) = \Omega_D^{(L+\tau)}$, and $\mathrm{Cov}(R^r_t, R^s_{t-\tau}) = \mathrm{Var}(G\sum_{l=0}^{L}(\sum_{j=0}^{L-l} F^j)V_{t+l}) = \Omega_D^{(L)}$. We conclude that $\mathrm{Var}(\tilde{R}^s) = \mathrm{Var}(R^s_{t-\tau} - R^r_t) = \Omega_D^{(L+\tau)} - \Omega_D^{(L)}$, and $\mathrm{Cov}(R^r, \tilde{R}^s) = \mathrm{Cov}(R^r_t, R^s_{t-\tau} - R^r_t) = 0$.

Proof of Proposition 8. First, recall that $\sigma_1^2 = \eta\mathrm{Var}(q_t)/[(\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2)\sum_{i=1}^T(\sigma_i^2/\sigma_1^2)]$. In the CFAR setting, $\mathrm{Var}(\hat{\epsilon}_t) = (1 - \eta)\mathrm{Var}(q_t)$, and $\mathrm{Var}(\hat{\delta}^E_{t,i}) = (\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2)\sigma_1^2(\sigma_i/\sigma_1)^2$. Therefore, by (15):

$$\Omega_D^{(L)} = \mathrm{Var}(q_t)\left\{\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l}\alpha^j\right)^2 - \eta\left[\sum_{l=0}^{L}\left(\sum_{j=0}^{L-l}\alpha^j\right)^2 - \frac{\sum_{i=1}^{\min(L,T)}(\sigma_i/\sigma_0)^2\sum_{l=0}^{L-i}\left(\sum_{j=0}^{L-i-l}\alpha^j\right)^2}{\sum_{i=1}^{T}(\sigma_i/\sigma_0)^2}\right]\right\}.$$

In the SMI setting, one easily shows that

$$\frac{\mathrm{Var}(\hat{\epsilon}_t)}{\mathrm{Var}(q_t)} = 1 - \eta\frac{(\rho\theta_r + \theta_s)^2}{\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2}, \quad \text{and}$$

$$\frac{\mathrm{Var}(\hat{\delta}^E_{t,i})}{\mathrm{Var}(q_t)} = \eta\frac{(\rho\theta_r + \theta_s)^2}{\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2}\cdot\frac{(\sigma_i/\sigma_0)^2}{\sum_{k=1}^{T}(\sigma_k/\sigma_0)^2}.$$

Hence, for the SMI setting, the value of $\Omega_D^{(L)}$ is given by the same expression as for the CFAR setting, but with η substituted with $\eta(\rho\theta_r + \theta_s)^2/(\theta_r^2 + 2\rho\theta_r\theta_s + \theta_s^2) \le \eta$. Next, let ψ_S and ψ_C be the values of $\sqrt{\Omega_D^{(L)}}$ under the SMI and CFAR settings, respectively. In view of the expressions above, we get $\psi_S - \psi_C = \sqrt{a - b_S\eta} - \sqrt{a - b_C\eta}$, for nonnegative values a, b_S, b_C, where $a \ge b_C \ge b_S$ (easy to see). The fact that the difference $\sqrt{a - b_S\eta} - \sqrt{a - b_C\eta}$ is increasing in η is verified by simple algebra. The proposition follows immediately.

References

Anand, K.S., H. Mendelson. 1997. Information and organization for horizontal multimarket coordination. *Management Sci.* **43** 1609–1627.

Andersen Consulting. 1998. Unlocking hidden value in the personal computer supply chain. Report.

Aviv, Y. 2001a. The effect of collaborative forecasting on supply chain performance. *Management Sci.* **47** (10) 1326–1343.

——. 2001b. A time-series framework for supply chain inventory management. Working paper, Washington University.

——, A. Federgruen. 1999. The operational benefits of information sharing and vendor managed inventory (VMI) programs. Working paper, Washington University.

Axsäter, S., K. Rosling. 1993. Notes: Installation vs. echelon stock policies for multilevel inventory control. *Management Sci.* **39** 1274–1280.

Cachon, G. P., M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Sci.* **46** (8) 1032–1048.

——, P. H. Zipkin. 1999. Competitive and cooperative inventory policies in a two-stage supply chain. *Management Sci.* **45** 936–953.

Chen, F. 1998. Echelon reorder points, installation reorder points, and the value of centralized demand information. *Management Sci.* **44** (12/2) 221–234.

——, Z. Drezner, J.K. Ryan, D. Simchi-Levi. 2000a. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Sci.* **46** (3) 436–443.

——, J.K. Ryan, D. Simchi-Levi. 2000b. The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Res. Logist.* **47** 269–286.

Cheung, K.L., H.L. Lee. 2002. The inventory benefit of shipment coordination and stock rebalancing in a supply chain. *Management Sci.* **48** (2) 300–306.

Clark, A.J., H. Scarf. 1960. Optimal policies for a multi-echelon inventory problem. *Management Sci.* **6** 475–490.

Croson R., Donohue, K.L. 2001. Experimental economics and supply chain management. Forthcoming, *Interfaces.*

Fiddis, C. 1997. Manufacturer-retailer relationships in the food and drink industry: Strategies and tactics in the battle for power. *FT Management Rep.,* Financial Times Retail & Consumer Publishing, London, U.K.

Frazier, R.M. 1986. Quick response in soft lines. *Discount Merchandiser* (Jan) 86–89.

Gavirneni, S., R. Kapuscinski, S. Tayur. 1999. Value of information in capacitated supply chains. *Management Sci.* **45** 16–24.

Graves, S. C. 1999. A single-item inventory model for a nonstationary demand process. *Manufacturing and Service Oper. Management* **1** 50–61.

Hammond, J. 1990. Quick response in the apparel industry. *Harvard Bus. School Note* 9-690-038.

Hayek, F.A. 1945. The use of knowledge in society. *Amer. Econom. Rev.* **35** 519–530.

Kahn, J. 1987. Inventories and the volatility of production. *Amer. Econom. Rev.* 667–679.

Koloczyc, G. 1998. Retailers, suppliers push joint sales forecasting. *Stores* (June).

Kurt Salomon Associates, Inc. 1993. *Efficient Consumer Response: Enhancing Consumer Value in The Grocery Industry.* Food Marketing Institute, Washington, D.C.

Lee H., S. Whang. 1999. Decentralized multi-echelon supply chains: Incentives and information. *Management Sci.* **45** 633–640.

——, ——. 2001. Demand chain excellence: A tail of two retailers. *Supply Chain Management Rev.* **5** 40–46.

——, K.C. So, C.S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Sci.* **46** 626–643.

Magretta, J., 1998. The power of virtual integration: An interview with Dell Computer's Michael Dell. *Harvard Bus. Rev.* (March–April) 73–84.

Mathews, R. 1995. Spartan pulls the plug on VMI. *Progressive Grocer* **74** (11).

Miyaoka, J., L.R. Kopczak. 2000. Using forecast synchronization to reduce the bullwhip effect of the forecast error. Working paper, Department of Management Science and Engineering, Stanford University, Stanford, CA.

Raghunathan, S. 2001. Information sharing in a supply chain: A note on its value when demand Is nonstationary. *Management Sci.* **47** 605–610.

Song, J.S., Zipkin, P. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41** 351–370.

Watson, N., Y.S. Zheng. 2000. Adverse effects of over-reaction to demand changes and improper demand forecasting. Working paper, University of Pennsylvania.

The consulting Senior Editor for this manuscript was Han Lee. This manuscript was received on December 30, 1999, and was with the authors 541 days for 4 revisions. The average review cycle time was 61 days.