

The Impact of External Demand Information on Parallel Supply Chains with Interacting Demand

Xiaolong Zhang¹

Yao Zhao²

Abstract

This paper considers two parallel supply chains with interacting demand streams. Each supply chain consists of one supplier and one retailer. The two demand streams are jointly described with a vector autoregressive time-series process in which they interact and their respective innovation errors correlate contemporaneously. For each supply chain, we develop insights into when and how much the supplier and the retailer can improve on their forecasting accuracy if the external demand history of the other supply chain is utilized. When this external demand history is not available or made available after a time lag, we develop a partial process and a delayed process to characterize the demand structure that the retailer can recover from the available demand histories. Our results show that the external demand history of the other supply chain always helps the retailer make better forecasts when demand streams interact; however, the enhanced information alters the retailer's order process, which may produce larger forecasting errors for the supplier. Conditions are established for the supplier to benefit from the external demand history of the other supply chain.

(Key Terms: Forecasting; Interacting Supply Chains; Inventory Control; Time-Series; Joint Demand Processes)

Original submission on 12/19/2006; 1st revision submitted on 3/21/2008; 2nd revision submitted on 3/3/2009; accepted on 8/25/2009.

¹ Department of Finance and Quantitative Analysis, College of Business Administration, Georgia Southern University, Statesboro, GA 30460-8151. Tel: (912) 478 5741, Fax: (912) 478 1835, Email: xzhang@georgiasouthern.edu.

² Department of Supply Chain Management and Marketing Science, Rutgers Business School – Newark and New Brunswick, Rutgers University, 180 University Avenue, Newark, NJ 07102. Email: yzhao@andromeda.rutgers.edu

The Impact of External Demand Information on Parallel Supply Chains with Interacting Demand

1. Introduction

Accurate demand forecasts are crucial to inventory planning and control. A knowledge of the expected demand during lead-time helps to set an appropriate level of anticipation stock, while measures of forecasting errors determine a proper level of safety stock. Within a company, the prevalence of scanner data has greatly reduced the cost of accessing demand data in real time for a variety of marketing and operations decisions (Bucklin and Gupta 1999). In addition, the advent of RFID and other Auto ID systems (McFarlane and Sheffi 2008) are expected to continue lowering the barrier of real time data acquisition. In order to navigate the business environment more efficiently, many companies are willing to pay a premium in acquiring business intelligence that includes competitors' sales data and pricing as well as overall market trends by brand, state and time. This type of business intelligence is provided by firms like BusinessObjects (an SAP company) (see <http://www.ondemand.com/information/>), which draws data from sources such as Thompson Financial and eBay Market Data. Their customers, companies using this data for operational decisions, are well documented in the press (see <http://www.ondemand.com/customers/default.asp>).

Motivated by these advances in practice, we attempt to address the following two questions in this paper:

(1) Under what conditions, if any, can a supply chain that acquires the retail level sales data of other supply chains improve its performance and demand forecast?

(2) How much improvement in forecast accuracy can be achieved?

So far, the value of the demand information in the context of multiple interacting (e.g., competing) demand streams has not been addressed in the literature of supply chain management. The multiple demand processes can be related in two distinct fashions. First, they may correlate stochastically, in which case we need to specify the nature of such correlation. Second, they may interact deterministically, in which case we can use a known deterministic function to describe such interaction. Neither case can be disposed of *a priori* because of market

forces. In the marketing literature, market effects among multiple demand streams (due to brand or product category) have been conveniently segregated into primary and secondary market effects. The primary market effect refers to the stochastic co-movements in demand processes caused by macro-economic dynamics and aggregate industrial marketing efforts. The secondary market effect represents the systematic interaction between demands caused by firm-level competitive efforts or product substitution (Schultz and Wittink 1976). The strength of the secondary market effect depends on market structure. For example, in a perfectly competitive market in which no company has a significant influence, we do not expect the demand streams of two companies to exhibit strong interactions. In a duopoly market, however, the demand streams faced by the two competitors are likely to have strong interactions whether they engage in a Nash-Cournot or a Bertrand competition (Varian 1992). As a result of these marketing and economic forces, changes in one demand stream can exert a predictable impact on the changes in another demand stream. Therefore, acquiring the information of a related demand stream may allow every party in a supply chain to significantly improve its forecasting accuracy.

In this paper, we consider two parallel supply chains where their external demand streams possess temporal dynamics driven by both the primary and the secondary market effects. The joint demand process is modeled by a VARMA (vector autoregressive and moving average) process. In particular, a VARMA process incorporates the primary market effect by allowing the errors in the joint process to correlate contemporaneously; it also takes the secondary market effect into account by linking the current demand of one supply chain to the lagged demands of the other. A detailed justification of this demand model is provided in Section 3. In the rest of the paper, we use *correlation* to represent the primary market effect and *interaction* to denote the secondary market effect in the joint demand process.

The objective of the paper is to develop insights into the value of acquiring information of the other demand stream. For this purpose, we focus on the following three scenarios: full-information, partial-information, and delayed-information forecasting. In the case of *full-information forecasting* (FIF), the joint demand processes are known to all parties in both supply chains. This information scenario may represent a centralized operational setting within a company in which demand information is shared across supply chains and easily accessible in real time. In the case of *partial-information forecasting* (PIF), all parties in each supply chain only know the external demand history of their own supply chain. PIF may represent parallel

supply chains that cross company boundaries, and information sharing is restricted due to confidentiality or lack of an effective demand information sharing mechanism. Given the limited information, the underlying joint demand process cannot be identified by all parties in each supply chain. We characterize the demand processes that we call the partial demand processes to describe what the retailers can recover and identify from their respective demand histories. In the case of *delayed-information forecasting* (DIF), we assume that the demand information of the other supply chain is made available after a time lag. This clearly represents a case in between FIF and PIF. We characterize the demand process identified from the available demand information for this case and derive the appropriate forecasting results accordingly.

We compare the forecasting errors of all parties involved among the three information settings of FIF, PIF and DIF. Among other results, our analytical and numerical studies show that FIF can substantially improve the forecast accuracies for all parties in a supply chain if there is a strong interaction between the demand processes of the parallel supply chains and demand autocorrelations are positive. However, when demands are interacting but demand autocorrelations are negative, the retailers' order processes under FIF have a larger variability than under PIF or DIF, which leads to larger forecast errors for the suppliers. This negative repercussion on suppliers is likely to occur in supply chains involving consumer products, many of which have a negative autocorrelation due to a substitution effect over time. Specifically, a high demand in one period tends to be followed by a low demand in the immediate future. Examples may include grocery items with a long shelf life such as canned food--a large purchase for a week reduces the need to buy in the following weeks, generating a negative demand autocorrelation.

1.1. Literature Review

Many authors have studied the benefits of acquiring the retail level demand information for the upstream parties in a serial supply chain. Lee, So, and Tang (2000) consider a two-stage serial supply chain in which the external demand follows AR(1) process and demonstrate that the inventory cost savings can be substantial if demand autocorrelation is highly positive. Raghunathan (2001) considers a similar model and shows that the external demand information becomes less valuable as the upstream supplier uses longer order history to forecast future orders. Miyaoka and Hausman (2004) consider an ARMA (0,1,1) demand process and show that

a stale forecast can reduce excessive demand fluctuations and thus the total supply chain cost. Gaur, Giloni, and Seshadri (2005) and Zhang (2004) extend Lee et al. (2000) to allow for a general ARMA demand processes and multi-echelon serial supply chains. Gilbert (2006) uses an even more encompassing class of ARIMA demand process to study the time-series structure of inventory level, orders, and order stability in a serial supply chain.

Gavirneni, Kapuscinski, and Tayur (1999) consider a two-stage serial supply chain in which the supplier has limited production capacity and receives POS data from the retailer at the end of each review period regardless of whether the retailer places orders or not. Assuming i.i.d demand distributions, the authors show that the cost savings from information sharing can be significant, especially when the production capacity is not tight. Simchi-Levi and Zhao (2003a) consider a two-stage system similar to Gavirneni et al. (1999), but assume that the supplier can receive POS data from the retailer even during a review period. The paper studies the impact of the frequency and timing of information sharing. Mishra, Raghunathan, and Yue (2009) investigate the value of sharing demand forecasts when the retailer in a serial supply chain can affect the retail demand by setting prices and the supplier can control the wholesale price charged to the retailer. They show that retailer's sharing of demand forecasts may sometimes hurt itself, and they design incentive compatible payment schemes to induce Pareto-optimal forecasts sharing. While most papers focus on the value of sharing downstream information, Choi, Blocher, and Gavirneni (2008) reverse the direction of sharing and consider sharing yield information from an upstream supplier to a downstream manufacturer.

For a distribution system with a single supplier and multiple retailers facing independent customer demands, Cachon and Fisher (2000) demonstrate that a 2.2% average reduction in inventory cost and a maximum reduction of 12.1% can be achieved by better allocating inventory among the retailers if the supplier has access to the retailers' demand information. In a similar model, Moinzadeh (2002) shows that the demand information can also help the supplier improve its replenishment policy and reduce the system-wide cost by an average of 3.2% and a maximum of 34.9%. Aviv and Fedegruen (1998) consider a similar model except that the suppliers' orders are subject to a capacity constraint. They report that information sharing can reduce system wide cost by up to 5%. Simchi-Levi and Zhao (2003b) consider single-supplier multi-retailer distribution systems where the retailers' demand processes are correlated but not

interacting. The authors quantified the impact of acquiring demand information from every retailer on the supplier's forecast accuracy.

In contrast to the previous work on supply chains with either a single demand stream or multiple non-interacting demand streams, this paper is a first attempt in the literature to study the value of acquiring demand information with multiple interacting demand streams. This situation can be found in either multiple interacting supply chains or multiple interacting products within one firm that are sourced from different suppliers.

Aviv (2003) considers a general state space model for the retailers' demand and develops a framework to assess the benefits of the information sharing agreements. His model can be shown to encompass the scenarios considered in this paper; however, the focus of Aviv (2003) is primarily on methodology. We instead focus on understanding how demand interaction affects the value of information sharing for various parties in the supply chain.

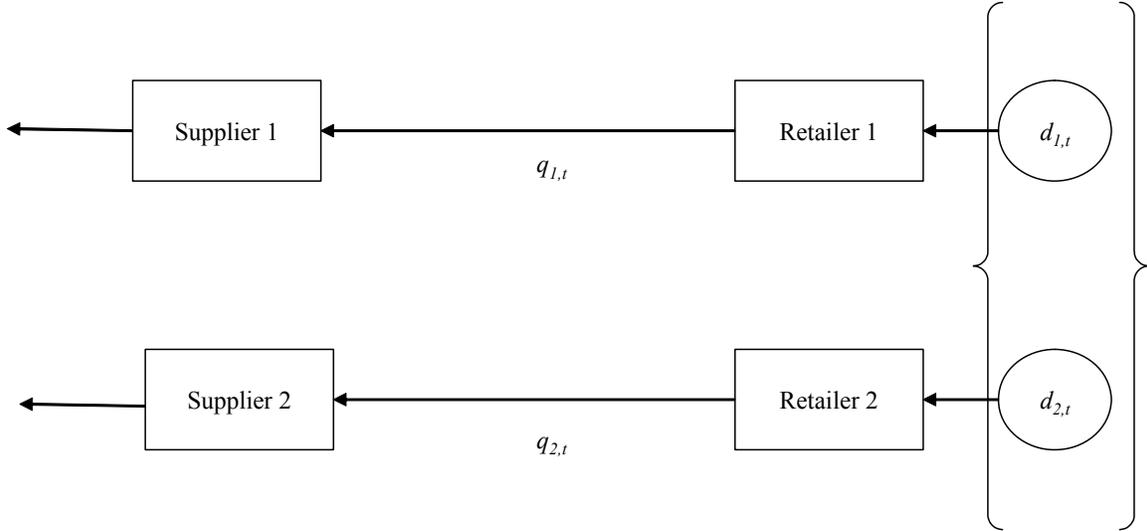
Our joint demand model represents a multivariate extension of the AR(1) model for quantifying the value of demand information sharing in a serial supply chain (Lee et al. 2000, Raghunathan 2001) and the bullwhip effect (Lee et al. 1997, Chen et al. 2000, Zhang 2005). It also extends the time-series demand modeling for studying various other issues in supply chains (Baganha and Cohen 1998, Graves 1999, Zhang 2004). Finally, the development of partial and delayed demand processes, to the best of our knowledge, is a new approach in understanding how different information settings affect the identifiable demand structures.

The rest of the paper is organized as follows. Section 2 describes the modeling framework and presents the joint demand process and its interpretation. Sections 3, 4, and 5 develop the forecasting results for all parties in each supply chain under FIF, PIF, and DIF respectively. Section 6 compares the forecasting accuracy among these information scenarios. Section 7 discusses the managerial implications of our results. Finally, Section 8 concludes the paper.

2. The Modeling Framework

2.1 Inventory Model

Consider two parallel supply chains where each consists of one supplier and one retailer. Each retailer faces an external demand stream. The following sketch depicts this system:



where

$d_{i,t}$: Demand observed by retailer $i = 1, 2$ during review period $[t-I, t]$.

$q_{i,t}$: Order placed by retailer $i = 1, 2$ at t after $d_{i,t}$ is realized but prior to the realization of $d_{i,t+1}$.

We follow the standard modeling assumptions in Lee et al. (2000), Aviv (2003) and Raghunathan (2001). In particular, we assume that all parties in the system control their inventories by periodic-review base-stock policies with a common review period (or ordering cycle, equivalently). For simplicity, we assume that the suppliers face zero replenishment lead-times, and the transportation lead-times between the retailers and their corresponding suppliers are also zero.

The sequence of events in a typical review period $[t, t+I]$ is as follows: at the beginning of this period after demand $d_{i,t}$ is realized, the retailers review their inventory levels and place orders to meet demand $d_{i,t+1}$. The suppliers receive the orders instantly and fill the orders from their on-hand inventories. We assume that demand not satisfied by the retailers is fully backordered. When the supplier experiences a stockout, we utilize the “borrow-return” assumption by Lee et al. (2000): the supplier can borrow (with certain penalty costs) from an external source with ample supply, and the borrowed stock is returned when inventory becomes available. This assumption is reasonable when the suppliers’ service levels are high. Finally, the suppliers update their forecasts.

Let $y_{i,t}$ and $y'_{i,t}$ represent the order-up-to levels of retailer i and supplier i at time t respectively. Under the assumption of normal demand distribution, the order-up-to levels are determined by (see Zipkin 2000, chapter 9):

$$y_{i,t} = \hat{d}_{i,t}(1) + k_i \sqrt{V(d_{i,t+1} - \hat{d}_{i,t}(1))}, \quad y'_{i,t} = \hat{q}_{i,t}(1) + k'_i \sqrt{V(q_{i,t+1} - \hat{q}_{i,t}(1))}, \quad \forall i = 1, 2, \quad (1)$$

where $\hat{d}_{i,t}(1)$ denotes the forecast made by retailer i at time t for the demand $d_{i,t+1}$, and $\hat{q}_{i,t}(1)$ denotes the forecast made by supplier i at time t for the retailer's order $q_{i,t+1}$ at time $t+1$. Terms associated with function $V(\cdot)$ represent the variances of lead-time demand forecast errors. Due to the zero lead-time assumption, the lead-time demand for the retailer refers to the demand during period t , and the lead-time demand for the supplier corresponds to the retailer's order made at beginning of the period $t+1$. In Eq. (1), k_i and k'_i are the normal z scores chosen to meet a desired level of stockout probability in a review period.

It is well known that the long-run average inventory holding and shortage costs are proportional to the standard deviation of the lead time demand forecast error (Zipkin 2000 and Aviv 2001), namely, $\sqrt{V(d_{i,t+1} - \hat{d}_{i,t}(1))}$ for the retailers and $\sqrt{V(q_{i,t+1} - \hat{q}_{i,t}(1))}$ for the suppliers. Therefore, a reduction in the variance of the lead-time demand forecast error can be directly translated into cost savings. For convenience, we exclude the stock acquisition (or purchasing) cost from our analysis. The exclusion does not affect our conclusions because this is a backorder model (see Zipkin 2000, pp. 415-420).

It follows from the assumptions of free return and full backlog at the retailers that the balance of flow yields the following relationship between order quantities and order-up-to levels:

$$q_{i,t} = y_{i,t} - y_{i,t-1} + d_{i,t}. \quad (2)$$

Substituting the order-up-to level from Eq. (1) in the above equation, we have

$$q_{i,t} = \left[\hat{d}_{i,t}(1) - \hat{d}_{i,t-1}(1) \right] + d_{i,t} + k_i \left[\sqrt{V(d_{i,t+1} - \hat{d}_{i,t}(1))} - \sqrt{V(d_{i,t} - \hat{d}_{i,t-1}(1))} \right].$$

where $\hat{d}_{i,t-1}(1)$ is the forecast made by retailer i at time $t-1$ for $d_{i,t}$. Because the identifiable demand processes in all three information settings of this paper are covariance stationary, the standard deviation of lead-time demand forecasting error is constant over time. Thus, the second term on the right hand side of the above equation vanishes, and we have

$$q_{i,t} = \left[\hat{d}_{i,t}(1) - \hat{d}_{i,t-1}(1) \right] + d_{i,t}. \quad (3)$$

2.2 The Joint Demand Model

The external demands faced by the retailers are modeled by the following first order vector autoregressive process:

$$\begin{aligned} d_{1,t} &= a_1 + \rho_{11}d_{1,t-1} + \rho_{12}d_{2,t-1} + \varepsilon_{1,t} \\ d_{2,t} &= a_2 + \rho_{21}d_{1,t-1} + \rho_{22}d_{2,t-1} + \varepsilon_{2,t}, \end{aligned} \quad (4)$$

where a_1 and a_2 are constant intercepts that determine the means of $d_{1,t}$ and $d_{2,t}$, ρ_{11} and ρ_{22} are the autocorrelation coefficients, ρ_{12} and ρ_{21} are the interaction coefficients, and the random error terms $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are jointly-distributed normal random variables with zero means and the following variance-covariance structure:

$$V(\varepsilon_{i,t}) = \sigma_{ii}^2, \quad COV(\varepsilon_{1,t}, \varepsilon_{2,t}) = \sigma_{12}, \quad COV(\varepsilon_{i,t}, \varepsilon_{j,t'}) = 0 \text{ if } t \neq t'. \quad (5)$$

The errors are independent across time but are correlated contemporaneously. The covariance between these errors represents the primary market effect caused by stochastic co-movements in the demand process. There are no prior restrictions on the sign of the autocorrelation and interaction coefficients as well as the covariance between the two errors. When $\rho_{12} = \rho_{21} = \sigma_{12} = 0$, the joint demand model reduces to two stochastically independent first-order autoregressive processes. As such, the joint demand model is a multivariate extension of the first-order autoregressive demand models. To satisfy *covariance stationarity*, the roots of the following characteristic equation must be inside the unit circle:³

$$(x - \rho_{11})(x - \rho_{22}) - \rho_{12}\rho_{21} = 0. \quad (6)$$

It is convenient to express the demand process in a mean-centered form in which the unconditional (stationary) means are subtracted from their demands. Denote μ_1 and μ_2 to be the unconditional mean demand for retailers 1 and 2. Taking unconditional expectation on both sides of the joint demand model, we have

$$\begin{aligned} \mu_1 &= a_1 + \rho_{11}\mu_1 + \rho_{12}\mu_2 \\ \mu_2 &= a_2 + \rho_{21}\mu_1 + \rho_{22}\mu_2. \end{aligned}$$

Solving the above equations obtains

$$\mu_1 = \frac{(1 - \rho_{22})a_1 + \rho_{12}a_2}{(1 - \rho_{11})(1 - \rho_{22}) - \rho_{12}\rho_{21}}, \quad \mu_2 = \frac{(1 - \rho_{11})a_2 + \rho_{21}a_1}{(1 - \rho_{11})(1 - \rho_{22}) - \rho_{12}\rho_{21}}.$$

³ See Hamilton (1994) for stationarity and invertibility conditions of a general VARMA process.

The denominator is always positive because of covariance-stationarity. To assure the positivity of unconditional expected demand, we further impose the following conditions:

$$a_1 > 0, \quad a_2 > 0, \quad (1 - \rho_{22}) + \rho_{12}a_2 > 0, \quad (1 - \rho_{11})a_2 + \rho_{21}a_1 > 0.$$

The joint demand model in Eq. (4) can now be rewritten as

$$\begin{aligned} z_{1,t} &= \rho_{11}z_{1,t-1} + \rho_{12}z_{2,t-1} + \varepsilon_{1,t} \\ z_{2,t} &= \rho_{21}z_{1,t-1} + \rho_{22}z_{2,t-1} + \varepsilon_{2,t}, \end{aligned} \tag{7}$$

where $z_{1,t} = d_{1,t} - \mu_1$, $z_{2,t} = d_{2,t} - \mu_2$. In the rest of the paper, we refer to $z_{i,t}$ as demand without causing any confusion.

The joint demand model dynamically links the current demand of a retailer to its own past demand and to the past demand of the other retailer. The interaction coefficients, ρ_{12} and ρ_{21} , capture the interaction between demands, while the autocorrelation coefficients, ρ_{11} and ρ_{22} , reflect the inertia and habit in consumer demand. A negative (positive) demand interaction is expected if the two supply chains are competitors (partners) or their products are substitutes (complements, respectively). Stockpiling of a retailer's product can also generate a negative (positive) demand interaction because it suppresses (boosts) the future demand for the other retailer if their products are substitutes (complements). Stockpiling can also generate a negative demand autocorrelation. For example, a week of high demand for Coke created by a deep discount will be followed by a lack of demand in the immediate future weeks. A positive demand autocorrelation is consistent with purchasing incidence models in the marketing research literature (Ehrenberg 1988). In particular, Morrison and Schmittlein (1988) analyze the changes in the number of purchases that a consumer makes over time, and they show that the expected number of purchases in the current period is a positive linear function of the number of previous purchases.

Using demand substitution and complementarity to explain demand interaction requires a subtle interpretation of our backorder assumption. When demand surpasses available stock in period t , the backlogged demand is not counted as part of the future d_{t+i} , $i > 0$, when the backorders are picked up. Instead, it is counted as part of the current d_t when the backorder occurs. This interpretation is consistent with most retailing practice. The backlogged sales are ticketed as demand when customers make the purchase, and pickups later are often not even recorded. As such, the forces of demand substitution and complementarity due to the primary and secondary market effects (e.g., changes in pricing, advertising and other economic factors from competing

firms) are assumed to act upon both satisfied demand and unsatisfied (or backlogged) demand in the exact same manner. These assumptions imply that in this paper we do not study demand substitution due to stockouts, where customers switch to a similar product when their original purchases are backordered, a phenomenon analyzed extensively in the revenue and assortment management literature (van Ryzin and Mahajan 1999). While the latter is more of an operational issue, the former is associated with market forces and thus is more of a strategic issue that aligns well with the objective of this paper.

The joint demand model is related to the Markovian brand choice models in the marketing literature (Lilien, et al., 1992, pp. 45-49), which involve a Markov chain with the choice of two retailers as the state. The transition matrix specifies the conditional probability of staying with the same retailer or switching to the other retailer, and the market share for each retailer evolves according to the following equation:

$$\begin{aligned} m_{1,t} &= p_{11}m_{1,t-1} + p_{21}m_{2,t-1} \\ m_{2,t} &= p_{12}m_{1,t-1} + p_{22}m_{2,t-1}. \end{aligned} \quad (8)$$

where p_{ij} is the probability of selecting retailer j in time t given that the customer chooses retailer i in time $t-1$, and $m_{i,t}$ is the market share for retailer i in time t . Let S be the fixed total market size and π_i be the stationary market share for each retailer that satisfies the following balance conditions:

$$\pi_1 = p_{11}\pi_1 + p_{21}\pi_2, \quad \pi_2 = p_{12}\pi_1 + p_{22}\pi_2.$$

Then the Markovian share evolution Eq. (8) can be rewritten in a deviation form such that

$$\begin{aligned} m_{1,t} - \pi_1 &= p_{11}(m_{1,t-1} - \pi_1) + p_{21}(m_{2,t-1} - \pi_2) \\ m_{2,t} - \pi_2 &= p_{12}(m_{1,t-1} - \pi_1) + p_{22}(m_{2,t-1} - \pi_2). \end{aligned} \quad (9)$$

Multiplying the equations in (9) by S , we see clearly that the Markovian brand choice model reproduces the dynamic portion of the joint demand model by $d_{i,t} = Sm_{i,t}$, $\mu_i = S\pi_i$ for $i = 1$ and 2 .

2.3 Parallel Supply Systems in Action and Demand Transparency

The parallel supply chain structure reflects the operational settings of many real-world supply chains. When residing within a single company, the structure may represent the parallel supply chains that either target at different geographical regions, or provide logistical support for different major brands, or serve different channels such as the traditional brick-mortar and online stores. When residing across different companies, the parallel structure may represent two com-

peting national retail chains, each supplied by a dedicated supplier. The two suppliers can be competitors as well, each running its own supply chain that fills orders for two products sold in the same store. In each of these cases, we clearly expect the demand streams for the two supply chains to be interacting and correlated.

To characterize demand information flow across the two serial chains, we can think of a transparency spectrum. At one extreme of the spectrum, high transparency implies that all demand data can be accessed in real time by all parties in the parallel system. At the other extreme of the spectrum, low transparency implies that each supply chain operates based upon only local demand information that does not flow across the two serial chains, although demand data is shared between the retailer and the supplier within each chain. In the middle of the spectrum, demand data sharing between the two supply chains is impeded in various degrees due to untimely data collection, data access, or other frictions in the supply chain.

The three information settings, FIF, PIF and DIF, considered in this paper reflect the high, low and medium degrees of transparency between the two supply chains respectively. In each setting, we assume that the retailer and the supplier in the same supply chain possess the same amount of demand information. This allows us to isolate the value of demand information from the other supply chain. The following table summarizes what the retailer and the supplier in each supply chain know about the demand in the three information settings:

PIF	DIF	FIF
Current demand history in its own supply chain	Information under PIF plus delayed demand history of the other supply chain	Current demand histories of both supply chains
Low Transparency		High Transparency

Demand Transparency Spectrum

3. Forecasting under Full-Information

In this section, we consider the case of *full-information* (FIF). We shall characterize the minimum mean square error (MMSE) forecasts for each party in a supply chain. For the definition and properties of MMSE forecast, we refer the reader to Box and Jenkins (1976).

We first derive the MMSE forecasts for the retailers. Let $\hat{z}_{i,t}(1)$ be the MMSE forecast for $z_{i,t+1}$ after $z_{i,t}$ is observed. The lead-time demand forecast and the associated forecasting error are given by

$$\hat{d}_{i,t}(1) = \mu_i + \hat{z}_{i,t}, \quad d_{i,t+1} - \hat{d}_{i,t}(1) = z_{i,t+1} - \hat{z}_{i,t}(1).$$

We use H_t to denote the set of information available for forecasting at time t , which consists of demands known to all parties in both supply chains up to and including t :

$$H_t = \{z_{1,t}, z_{2,t}, z_{1,t-1}, z_{2,t-1}, \dots\}. \quad (10)$$

Based upon standard time-series results (Box and Jenkins 1976), the MMSE forecast $\hat{z}_{i,t}(1)$ is simply the conditional expectation of $z_{i,t+1}$ given H_t :

$$\hat{z}_{i,t}(1) = E(z_{i,t+1} | H_t).$$

From the joint demand model, the future demands $z_{1,t+1}$ and $z_{2,t+1}$ can be decomposed as a linear function of most recently observed demands $z_{1,t}$ and $z_{2,t}$ and future unobserved errors $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$,

$$z_{1,t+1} = \rho_{11}z_{1,t} + \rho_{12}z_{2,t} + \varepsilon_{1,t+1}, \quad z_{2,t+1} = \rho_{21}z_{1,t} + \rho_{22}z_{2,t} + \varepsilon_{2,t+1}. \quad (11)$$

The fact that the future errors $\varepsilon_{1,t+1}$ and $\varepsilon_{2,t+1}$ are independent of the observed demands implies that in the case of FIF, the retailers' MMSE forecasts are given by

$$\hat{z}_{1,t}(1) = \rho_{11}z_{1,t} + \rho_{12}z_{2,t}, \quad \hat{z}_{2,t}(1) = \rho_{21}z_{1,t} + \rho_{22}z_{2,t}. \quad (12)$$

Their forecasting errors and the associated variances are given as follows:

$$d_{i,t+1} - \hat{d}_{i,t}(1) = \varepsilon_{i,t+1}, \quad V[d_{i,t+1} - \hat{d}_{i,t}(1)] = \sigma_{ii}^2, \quad \forall i = 1, 2. \quad (13)$$

We next derive $\hat{q}_{i,t}(1)$, the MMSE forecasts for the suppliers. Substituting the retailers' lead-time demand forecasts from Eq. (12) into Eq. (3), we obtain the order quantities for period t as

$$\begin{aligned} q_{1,t} &= \hat{d}_{1,t}(1) - \hat{d}_{1,t-1}(1) + d_{1,t} = \rho_{11}(z_{1,t} - z_{1,t-1}) + \rho_{12}(z_{2,t} - z_{2,t-1}) + z_{1,t} + \mu_1 \\ q_{2,t} &= \hat{d}_{2,t}(1) - \hat{d}_{2,t-1}(1) + d_{2,t} = \rho_{21}(z_{1,t} - z_{1,t-1}) + \rho_{22}(z_{2,t} - z_{2,t-1}) + z_{2,t} + \mu_2. \end{aligned} \quad (14)$$

These two orders placed by the retailers are the suppliers' demands, and the above equations link suppliers' demand to the external demand that the retailers face. The sequence of events in an ordering cycle implies that the suppliers' ordering decisions are made after receiving orders from their respective retailers. At that time, the suppliers have all the joint demand history H_t avail-

able to forecast the retailers' next order. Therefore, the MMSE forecast $\hat{q}_{i,t}(1)$ of supplier i for $q_{i,t+1}$ is again the conditional expectation of $q_{i,t+1}$ given H_t , i.e.,

$$\hat{q}_{i,t}(1) = E(q_{i,t+1} | H_t).$$

It follows from Eq. (14) that,

$$\begin{aligned} q_{1,t+1} &= \hat{d}_{1,t+1}(1) - \hat{d}_{1,t}(1) + d_{1,t+1} = \rho_{11}(z_{1,t+1} - z_{1,t}) + \rho_{12}(z_{2,t+1} - z_{2,t}) + z_{1,t+1} + \mu_1 \\ q_{2,t+1} &= \hat{d}_{2,t+1}(1) - \hat{d}_{2,t}(1) + d_{2,t+1} = \rho_{21}(z_{1,t+1} - z_{1,t}) + \rho_{22}(z_{2,t+1} - z_{2,t}) + z_{2,t+1} + \mu_2. \end{aligned}$$

Utilizing the joint demand process in Eq. (7), we obtain the following expressions:

$$\begin{aligned} q_{1,t+1} &= (\rho_{11}^2 + \rho_{12}\rho_{21})z_{1,t} + \rho_{12}(\rho_{11} + \rho_{22})z_{2,t} + (1 + \rho_{11})\varepsilon_{1,t+1} + \rho_{12}\varepsilon_{2,t+1} + \mu_1 \\ q_{2,t+1} &= \rho_{21}(\rho_{22} + \rho_{11})z_{1,t} + (\rho_{22}^2 + \rho_{12}\rho_{21})z_{2,t} + (1 + \rho_{22})\varepsilon_{2,t+1} + \rho_{21}\varepsilon_{1,t+1} + \mu_2. \end{aligned} \quad (15)$$

The decomposition immediately yields the forecasting results for the suppliers in the case of FIF.

The suppliers' MMSE forecasts for the retailers' order quantities are given by

$$\begin{aligned} \hat{q}_{1,t}(1) &= (\rho_{11}^2 + \rho_{12}\rho_{21})z_{1,t} + \rho_{12}(\rho_{11} + \rho_{22})z_{2,t} + \mu_1 \\ \hat{q}_{2,t}(1) &= \rho_{21}(\rho_{22} + \rho_{11})z_{1,t} + (\rho_{22}^2 + \rho_{12}\rho_{21})z_{2,t} + \mu_2, \end{aligned} \quad (16)$$

and the suppliers' forecasting errors satisfy

$$\begin{aligned} q_{1,t+1} - \hat{q}_{1,t}(1) &= (1 + \rho_{11})\varepsilon_{1,t+1} + \rho_{12}\varepsilon_{2,t+1}, & q_{2,t+1} - \hat{q}_{2,t}(1) &= (1 + \rho_{22})\varepsilon_{2,t+1} + \rho_{21}\varepsilon_{1,t+1} \\ V[q_{1,t+1} - \hat{q}_{1,t}(1)] &= (1 + \rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12} \\ V[q_{2,t+1} - \hat{q}_{2,t}(1)] &= (1 + \rho_{22})^2 \sigma_{22}^2 + \rho_{21}^2 \sigma_{11}^2 + 2(1 + \rho_{22})\rho_{21}\sigma_{12}. \end{aligned} \quad (17)$$

4. Forecasting under Partial-Information

In this section, we consider the case of *partial-information* (PIF) where the retailer and supplier in each supply chain only know their own external demand history. As a result, the joint demand structure is only *partially* known to each retailer, and the demand process that each supply chain can recover has a different autocorrelation structure than the joint demand process, which we call a partial demand process. In the rest of this section, we derive this partial demand process and determine the resulting MMSE forecasts for each party in the supply chains. For simplification, we use $z_{i,t}$ to represent the observed demand histories under both PIF and FIF. The context under which $z_{i,t}$ appears makes it clear which information setting is relevant, so no confusion arises.

4.1 Partial Demand Process

The demand information available for retailer i to make a forecast at time t is given by

$$H_{i,t} = \{z_{i,t}, z_{i,t-1}, \dots\}, \quad \forall i = 1, 2, \quad (18)$$

In this case, the retailer cannot identify the cross-interaction structure in the joint demand process. However, it can recover the demand autocorrelation structure from its observed demand and, at best, identify a partial demand process that is stationary and invertible.

To derive the partial processes, we first obtain a reduced form of the joint demand process as follows. First, take a one-period lag of $z_{2,t}$ in the 2nd line of the joint demand in Eq. (7) and substitute the resulting $z_{2,t-1}$ into the right hand side of $z_{1,t}$. We have

$$z_{1,t} = \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}z_{1,t-2} + \rho_{12}\rho_{22}z_{2,t-2} + \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1}. \quad (19)$$

We can also lag $z_{1,t}$ in Eq. (7) to obtain $z_{1,t-1}$ and multiply both sides of $z_{1,t-1}$ with ρ_{22} and arrive at

$$\rho_{22}z_{1,t-1} = \rho_{11}\rho_{22}z_{1,t-2} + \rho_{12}\rho_{22}z_{2,t-2} + \rho_{22}\varepsilon_{1,t-1}. \quad (20)$$

If we subtract Eq. (20) from Eq. (19), we can cancel $z_{2,t-2}$, and the resulting expression after re-grouping terms yields the following alternative demand process for Retailer 1:

$$z_{1,t} = \beta_1 z_{1,t-1} + \beta_2 z_{1,t-2} + v_{1,t}, \quad (21)$$

where $\beta_1 = \rho_{11} + \rho_{22}$, $\beta_2 = \rho_{12}\rho_{21} - \rho_{11}\rho_{22}$, and $v_{1,t} = \varepsilon_{1,t} - \rho_{22}\varepsilon_{1,t-1} + \rho_{12}\varepsilon_{2,t-1}$. By interchanging the role of $z_{1,t}$ and $z_{2,t}$ in the above derivation, we can obtain a similar process for Retailer 2 as follows:

$$z_{2,t} = \beta_1 z_{2,t-1} + \beta_2 z_{2,t-2} + v_{2,t}, \quad (22)$$

where $v_{2,t} = \varepsilon_{2,t} - \rho_{11}\varepsilon_{2,t-1} + \rho_{21}\varepsilon_{1,t-1}$.⁴ When the retailers only know their own demand history, they can identify the autoregressive components in Eq. (21) and Eq. (22) via a standard Box-Jenkins time-series approach. In addition, they can also identify the autocorrelation structures of the errors $v_{1,t}$ and $v_{2,t}$, but they cannot identify the cross-correlation structure $v_{1,t}$ and $v_{2,t}$. In addition, they lose the capability of separating out their individual terms $\varepsilon_{1,t}$, $\varepsilon_{1,t-1}$, $\varepsilon_{2,t}$, and $\varepsilon_{2,t-1}$. As a result, retailers cannot recover the joint demand process completely from their own demand histories. Theorem 1 provides us with the structure of the demand process identifiable under PIF when the underlying joint demand process is governed by the joint VARMA process.

⁴ In time-series literature, the reduced forms of the joint demand in Equations (21) and (22) correspond to what are commonly known as the marginal processes (Wei 1990, p.340).

Theorem 1: The demand processes that retailer i can recover based on its own demand history can be represented as a stationary and invertible ARMA(2, 1) process as follows:

$$\begin{aligned} z_{i,t} &= \beta_1 z_{i,t-1} + \beta_2 z_{i,t-2} + \xi_{i,t} - \theta_i \xi_{i,t-1}, \quad \forall i = 1, 2 \\ \theta_i &= \frac{-V(v_{i,t}) + \sqrt{[V(v_{i,t})]^2 - 4[COV(v_{i,t}, v_{i,t-1})]^2}}{2COV(v_{i,t}, v_{i,t-1})}, \end{aligned} \quad (23)$$

where $|\theta_i| < 1$, and θ_i is real. The error terms have zero means with the following variance and covariance structure:

$$E(\xi_{i,t} \xi_{i,t'}) = 0, \quad \forall t \neq t', \quad E(\xi_{i,t}^2) = \sigma_{i\xi}^2 = V(v_{ii}) / (1 + \theta_i^2).$$

Proof: See the appendix.

Theorem 1 indicates that the retailer's partial demand process can be considered as a filtered signal of the joint demand when demand information is restricted. The information setting in this case alters the demand processes that the retailers perceive, an observation that has not been recorded in the literature. The existence of the partial demand processes facilitates the derivation of the forecasting results under PIF.

The retailers' partial demand processes share a common autoregressive structure as the reduced forms in Eqs. (20)-(21), and the moving average component has the same order but with a distinct coefficient. It is worth noting that each of the errors $\xi_{1,t}$ and $\xi_{2,t}$ in the partial processes is temporally independent. However, they are cross-correlated at all lags because:

$$\begin{aligned} \xi_{1,t} - \theta_1 \xi_{1,t-1} &= \varepsilon_{1,t} - \rho_{22} \varepsilon_{1,t-1} + \rho_{12} \varepsilon_{2,t-1} \\ \xi_{2,t} - \theta_2 \xi_{2,t-1} &= \varepsilon_{2,t} - \rho_{11} \varepsilon_{1,t-1} + \rho_{21} \varepsilon_{1,t-1}. \end{aligned} \quad (24)$$

From their own demand histories $H_{1,t} = \{z_{1,t}, z_{1,t-1}, \dots\}$ and $H_{2,t} = \{z_{2,t}, z_{2,t-1}, \dots\}$, both retailers can identify and estimate their respective partial demand processes. Given that they are invertible, retailers can recover the error histories $\{\xi_{1,t}, \xi_{1,t-1}, \dots\}$ and $\{\xi_{2,t}, \xi_{2,t-1}, \dots\}$ from their corresponding demand histories. Therefore, knowing the demand histories is equivalent to knowing the error histories $\{\xi_{1,t}, \xi_{1,t-1}, \dots\}$ and $\{\xi_{2,t}, \xi_{2,t-1}, \dots\}$. However, the retailers are incapable of further identifying the errors $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ of the joint demand process that make up the errors $\xi_{1,t}$ and $\xi_{2,t}$. In the

rest of the paper, we call $\xi_{1,t}$ and $\xi_{2,t}$ partial errors to distinguish them from the joint demand errors $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$.

4.2 Forecasting with Partial Demand Process

Each retailer now makes lead-time demand forecasts based on its partial ARMA(2, 1) demand process. Unlike the case of FIF, the MMSE forecast $\hat{z}_{i,t}(1)$ is the conditional expectation with respect to $H_{1,t}$ instead of the complete demand history H_t :

$$\hat{z}_{i,t}(1) = E(z_{i,t+1} | H_{1,t}).$$

If we substitute backwards the partial demand process recursively, we can establish that the future demand for each retailer is a sum of its observed demands, past partial errors, and unobserved future partial errors. For example, the immediate future demands based on the partial processes can be expressed as follows:

$$\begin{aligned} z_{1,t+1} &= \beta_1 z_{1,t} + \beta_2 z_{1,t-1} + \xi_{1,t+1} - \theta_1 \xi_{1,t} \\ z_{2,t+1} &= \beta_1 z_{2,t} + \beta_2 z_{2,t-1} + \xi_{2,t+1} - \theta_2 \xi_{2,t}. \end{aligned}$$

Because the partial errors $\xi_{1,t}$ and $\xi_{2,t}$ are temporally independent, the one-period-ahead forecasts and forecasting errors are simply given by

$$\begin{aligned} \hat{z}_{i,t}(1) &= E(z_{i,t+1} | H_{i,t}) = \beta_1 z_{i,t} + \beta_2 z_{i,t-1} - \theta_i \xi_{i,t} \\ d_{i,t+1} - \hat{d}_{i,t}(1) &= z_{i,t+1} - \hat{z}_{i,t}(1) = \xi_{i,t+1} \quad \forall i = 1, 2. \end{aligned} \tag{25}$$

The observed partial errors $\xi_{1,t}$ and $\xi_{2,t}$ can be calculated from the observed demands because the partial processes are invertible (see Box and Jenkins, 1976, for a procedure). The forecasting errors immediately lead to the retailers' lead-time forecasting error variances under PIF:

$$V[d_{i,t+1} - \hat{d}_{i,t}(1)] = \sigma_{i\xi}^2, \quad \forall i = 1, 2. \tag{26}$$

Substituting the lead-time demand forecasts in Eq. (25) and the partial demand processes into Eq. (3), we obtain the following expression for the order quantities under PIF:

$$\begin{aligned} q_{i,t+1} &= \hat{d}_{i,t+1}(1) - \hat{d}_{i,t}(1) + d_{i,t+1} \\ &= \hat{d}_{i,t+1}(1) + \xi_{i,t+1} \\ &= (\beta_1^2 + \beta_2) z_{i,t} + \beta_1 \beta_2 z_{i,t-1} + \beta_1 \theta_i \xi_{i,t} + (1 + \beta_1 - \theta_i) \xi_{i,t+1} + \mu_i. \end{aligned}$$

The above equation decomposes the order quantities into the observed demand history and the future partial error. Consequently, the suppliers' forecasts and forecasting errors under PIF are given by:

$$\begin{aligned}\hat{q}_{i,t}(1) &= (\beta_1^2 + \beta_2)z_{i,t} + \beta_1\beta_2z_{i,t-1} + \beta_1\theta_i\xi_{i,t} + \mu_i \\ q_{i,t+1} - \hat{q}_{i,t}(1) &= (1 + \beta_1 - \theta_i)\xi_{i,t+1},\end{aligned}$$

which immediately yields the suppliers' forecasting error variances:

$$V[q_{i,t+1} - \hat{q}_{i,t}(1)] = (1 + \beta_1 - \theta_i)^2 \sigma_{i\xi}^2. \quad (27)$$

5. Forecasting under Delayed-Information

In this section, we consider the case of *delayed-information* (DIF) where retailer i ($= 1, 2$) knows its own external demand history and also the demand history of retailer j ($j=1, 2$ and $j \neq i$) with a delay of l_i periods. We will first derive the demand structure identifiable under DIF and then use it to obtain forecasting results for the retailers and suppliers.

The demand information available for retailer i to make a forecast at time t is given by

$$\tilde{H}_{i,t} = \{z_{i,t}, z_{i,t-1}, \dots, z_{j,t-l_i}, z_{j,t-l_i-1}, \dots\}, \quad \forall i, j = 1, 2 \text{ and } i \neq j. \quad (28)$$

We use a tilde, “ \sim ,” to differentiate the delayed information history under DIF from the partial demand history under PIF. As we would expect, retailer i still cannot identify the true joint demand process with the delayed demand history of the other supply chain, but it can do better than the marginal process. To derive this delayed demand process, we can back substitute $z_{2,t}$ from the joint demand into $z_{1,t}$ l_1 times in the following manner:

$$\begin{aligned}z_{1,t} &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}z_{1,t-2} + \rho_{12}\rho_{22}z_{2,t-2} + \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1} \\ &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}z_{1,t-2} + \rho_{12}\rho_{22}(\rho_{21}z_{1,t-3} + \rho_{22}z_{2,t-3} + \varepsilon_{2,t-2}) + \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1} \\ &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}z_{1,t-2} + \rho_{12}\rho_{21}\rho_{22}z_{1,t-3} + \rho_{12}\rho_{22}^2z_{2,t-3} \\ &\quad + \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1} + \rho_{12}\rho_{22}\varepsilon_{2,t-2} \\ &\quad \vdots \\ &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}(z_{1,t-2} + \rho_{22}z_{1,t-3} + \dots + \rho_{22}^{l_1-1}z_{1,t-l_1-1}) + \rho_{12}\rho_{22}^{l_1}z_{2,t-l_1-1} \\ &\quad + \varepsilon_{1,t} + \rho_{12}(\varepsilon_{2,t-1} + \rho_{22}\varepsilon_{2,t-2} + \dots + \rho_{22}^{l_1-1}\varepsilon_{2,t-l_1}).\end{aligned} \quad (29)$$

By symmetry, we can also obtain

$$\begin{aligned}
z_{2,t} &= \rho_{22}z_{2,t-1} + \rho_{12}\rho_{21}\left(z_{2,t-2} + \rho_{11}z_{2,t-3} + \cdots + \rho_{11}^{l_2-1}z_{2,t-l_2-1}\right) + \rho_{21}\rho_{11}^{l_2}z_{1,t-l_2-1} \\
&+ \varepsilon_{2,t} + \rho_{21}\left(\varepsilon_{1,t-1} + \rho_{11}\varepsilon_{1,t-2} + \cdots + \rho_{11}^{l_2-1}\varepsilon_{1,t-l_2}\right).
\end{aligned} \tag{30}$$

These two demand processes under DIF allow us to obtain the needed forecasting results for all parties.

Theorem 2: *The demand process under DIF converges to the partial demand process when $l_i \rightarrow \infty$.*

Proof: See the appendix.

Theorem 2 is not surprising. With an infinite delay, each retailer operates as if it only has its own demand history available. This reflects the intuition that DIF represents a compromise between PIF and DIF. Let

$$\tilde{v}_{1,t} = \varepsilon_{1,t} + \rho_{12}\left(\varepsilon_{2,t-1} + \rho_{22}\varepsilon_{2,t-2} + \cdots + \rho_{22}^{l_1-1}\varepsilon_{2,t-l_1}\right)$$

and

$$\tilde{v}_{2,t} = \varepsilon_{2,t} + \rho_{21}\left(\varepsilon_{1,t-1} + \rho_{11}\varepsilon_{1,t-2} + \cdots + \rho_{11}^{l_2-1}\varepsilon_{1,t-l_2}\right)$$

be the respective errors associated with the demand processes under DIF in (29) and (30). From its delayed demand history $\tilde{H}_{1,t}$, Retailer 1 can no longer uncover the individual joint errors $\varepsilon_{2,t-1}$, $\varepsilon_{2,t-2}$, and $\varepsilon_{2,t-l_1}$, which brings additional unpredictability beyond the FIF scenario. It is not an easy task, however, to determine exactly what each retailer can parse out from its own demand history and the delayed demand history of its counterpart. Fortunately, we can show that $\{z_{i,t}, z_{j,t-l_i}\} \forall i, j = 1, 2$ and $i \neq j$, namely, the demand of retailer i and the delayed demand of retailer j , possess a joint VARMA structure that can be used to determine the MMSE forecast for retailer i , conditioning on $\tilde{H}_{i,t}$.

Theorem 3: $\{z_{i,t}, z_{j,t-l_i}\} \forall i, j = 1, 2$ and $i \neq j$ can be represented as a joint VARMA(l_i, l_i) process.

Let $\gamma_1(0) = V(\tilde{v}_{1,t})$ be the unconditional variance of error term $\tilde{v}_{1,t}$ and $\gamma_1(k) = COV(\tilde{v}_{1,t}, \tilde{v}_{1,t-k})$ be the autocovariance of $\tilde{v}_{1,t}$ at lag k . For Retailer 1, we have

$$\begin{aligned}
\begin{bmatrix} z_{1,t} \\ z_{2,t-l_1} \end{bmatrix} &= \begin{bmatrix} \rho_{11} & \rho_{12}\rho_{22}^{l_1} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-l_1-1} \end{bmatrix} + \begin{bmatrix} \rho_{12}\rho_{21} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-2} \\ z_{2,t-l_1-2} \end{bmatrix} + \begin{bmatrix} \rho_{12}\rho_{21}\rho_{22} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-3} \\ z_{2,t-l_1-3} \end{bmatrix} \\
&+ \dots + \begin{bmatrix} \rho_{12}\rho_{21}\rho_{22}^{l_1-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t-l_1-1} \\ z_{2,t-2l_1-1} \end{bmatrix} \\
&+ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1,t} \\ \tilde{\varepsilon}_{2,t} \end{bmatrix} + \begin{bmatrix} \tilde{\theta}_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1,t-1} \\ \tilde{\varepsilon}_{2,t-1} \end{bmatrix} + \dots + \begin{bmatrix} \tilde{\theta}_{1,l_1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1,t-l_1} \\ \tilde{\varepsilon}_{2,t-l_1} \end{bmatrix}
\end{aligned} \tag{31}$$

where

$$\begin{aligned}
\tilde{\varepsilon}_{2,t} &= \varepsilon_{2,t-l_1}, \quad V(\tilde{\varepsilon}_{1,t}) = \tilde{\sigma}_{11}^2, \quad COV(\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t}) = \rho_{22}^{l_1-1} \sigma_{12} \\
COV(\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{j,t'}) &= 0 \text{ if } t \neq t'.
\end{aligned}$$

and the moving average coefficients $\tilde{\theta}_{1,1}$ to $\tilde{\theta}_{1,l_1}$ and $V(\tilde{\varepsilon}_{1,t}) = \tilde{\sigma}_{11}^2$ are obtained from solving the following nonlinear system of equations:

$$\begin{cases} (1 + \tilde{\theta}_{1,1}^2 + \tilde{\theta}_{1,2}^2 + \dots + \tilde{\theta}_{1,l_1}^2) \tilde{\sigma}_{11}^2 = \gamma_1(0) \\ (-\tilde{\theta}_{1,k} + \tilde{\theta}_{1,1}\tilde{\theta}_{1,k+1} + \dots + \tilde{\theta}_{1,l_1-k}\tilde{\theta}_{1,l_1}) \tilde{\sigma}_{11}^2 = \gamma_1(k), \quad k = 1, 2, \dots, l_1. \end{cases} \tag{32}$$

Note that we adopted the convention $\tilde{\theta}_{1,0} = 0$. For Retailer 2, the same result can be established with the obvious modifications.

Proof: See the appendix.

The advantage of developing the VARMA representation for $\{z_{i,t}, z_{j,t-l_i}\} \forall i, j = 1, 2$, and $i \neq j$ is that we know exactly what the retailers can recover from their information set $\tilde{H}_{i,t}$. Knowing $\tilde{H}_{i,t}$ is equivalent to knowing

$$\{\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{i,t-1}, \dots, \tilde{\varepsilon}_{j,t}, \tilde{\varepsilon}_{j,t-1}, \dots\} = \{\varepsilon_{i,t}, \varepsilon_{i,t-1}, \dots, \varepsilon_{j,t-l_i}, \varepsilon_{j,t-l_i-1}, \dots\}.$$

Using the joint demand that are lagged l_i periods, retailer i can also recover its own delayed errors $\{\varepsilon_{i,t-l_i}, \varepsilon_{i,t-l_i-1}, \dots\}$. The retailers, however, cannot separate out the joint errors $\{\varepsilon_{i,t}, \varepsilon_{i,t-1}, \dots, \varepsilon_{i,t-l_i+1}\}$ from delayed errors $\{\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{i,t-1}, \dots, \tilde{\varepsilon}_{i,t-l_i+1}\}$ that occurred after period $t-l_i$. The delayed errors $\{\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{i,t-1}, \dots, \tilde{\varepsilon}_{i,t-l_i+1}\}$ are filtered signals that mix the joint errors $\{\varepsilon_{i,t}, \varepsilon_{i,t-1}, \dots, \varepsilon_{i,t-l_i+1}\}$ in complex ways as described in Theorem 3.

The VARMA representation for $\{z_{i,t}, z_{j,t-l_i}\} \forall i, j = 1, 2$, and $i \neq j$ has a more important advantage. It facilitates deriving the MMSE forecasting result for both retailers and suppliers. Without this representation, we can employ Eq. (29) to derive the MMSE forecasts for Retailer 1. The difficulty here is that Retailer 2's demand errors after period $t - l_i$, namely, $\varepsilon_{2,t-k}, \forall 1 \leq k < l_1$, are correlated with the demand observed later. We need to forecast backwards these errors conditioning on the observed demand for Retailer 1. Except for short delays $l_i \leq 2$, analytical results on back forecasting are extremely messy, and we develop the VARMA representation to deal with this issue.

Retailers' forecasts and forecasting errors are now easily obtained from the DIF VARMA demand process. We have for $\forall i, j = 1, 2$, and $i \neq j$

$$\begin{aligned} \hat{z}_{i,t}(1) &= E(z_{i,t+1} | \tilde{H}_{i,t}) = \rho_{ii}z_{i,t} + \rho_{ij}\rho_{ji} \left(z_{i,t-1} + \rho_{jj}z_{i,t-2} + \dots + \rho_{jj}^{l_i-1}z_{i,t-l_i} \right) + \rho_{ij}\rho_{jj}^{l_i}z_{j,t-l_i} \\ &\quad - \tilde{\theta}_{i,1}\tilde{\varepsilon}_{i,t} - \tilde{\theta}_{i,2}\tilde{\varepsilon}_{i,t-1} - \dots - \tilde{\theta}_{i,l_i}\tilde{\varepsilon}_{i,t+1-l_i} \\ d_{i,t+1} - \hat{d}_{i,t}(1) &= z_{i,t+1} - \hat{z}_{i,t}(1) = \tilde{\varepsilon}_{i,t+1}. \end{aligned} \quad (33)$$

where $\tilde{\varepsilon}_{i,t-k}$ for $k \geq 1$ are observed errors. The forecasting errors immediately lead to the retailers' lead-time forecasting error variances under DIF:

$$V[d_{i,t+1} - \hat{d}_{i,t}(1)] = \tilde{\sigma}_{ii}^2, \forall i = 1, 2. \quad (34)$$

Substituting the lead-time demand forecasts in Eq. (33) and the partial demand processes into Eq. (3), we obtain the following expression for the order quantities under PIF:

$$\begin{aligned} q_{i,t+1} &= \hat{d}_{i,t+1}(1) - \hat{d}_{i,t}(1) + d_{i,t+1} \\ &= \rho_{ii}z_{i,t+1} + \rho_{ij}\rho_{ji} \left(z_{i,t} + \rho_{jj}z_{i,t-1} + \dots + \rho_{jj}^{l_i-1}z_{i,t+1-l_i} \right) + \rho_{ij}\rho_{jj}^{l_i}z_{j,t+1-l_i} \\ &\quad + \tilde{\varepsilon}_{i,t+1} - \tilde{\theta}_{i,1}\tilde{\varepsilon}_{i,t+1} - \tilde{\theta}_{i,2}\tilde{\varepsilon}_{i,t} - \dots - \tilde{\theta}_{i,l_i}\tilde{\varepsilon}_{i,t+2-l_i}. \end{aligned}$$

Given the observed demand history, only the first term after the second equal sign in the above expression involves future errors implicitly. Consequently, we can use this decomposition to derive the suppliers' forecasts and forecasting errors under DIF as follows

$$\begin{aligned} \hat{q}_{i,t}(1) &= E(q_{i,t+1} | \tilde{H}_{i,t}) \\ &= \rho_{ii}\hat{z}_{i,t}(1) + \rho_{ij}\rho_{ji} \left(z_{i,t} + \rho_{jj}z_{i,t-1} + \dots + \rho_{jj}^{l_i-1}z_{i,t+1-l_i} \right) + \rho_{ij}\rho_{jj}^{l_i}z_{j,t+1-l_i} \\ &\quad - \tilde{\theta}_{i,2}\tilde{\varepsilon}_{i,t} - \dots - \tilde{\theta}_{i,l_i}\tilde{\varepsilon}_{i,t+2-l_i}. \\ q_{i,t+1} - \hat{q}_{i,t}(1) &= (1 + \rho_{ii} - \tilde{\theta}_{i,1})\tilde{\varepsilon}_{i,t+1}, \end{aligned}$$

which immediately yields the suppliers' forecasting error variances:

$$V[q_{i,t+1} - \hat{q}_{i,t}(1)] = (1 + \rho_{ii} - \tilde{\theta}_{i,1})^2 \tilde{\sigma}_{ii}^2. \quad (35)$$

6. Supply Chain Performances under FIF, PIF and DIF

In this section, we compare the forecasting accuracies of the retailer and supplier in each supply chain among the three information settings. The comparisons provide analytical insights into the value of external demand information of the other supply chain.

6.1. FIF vs. PIF

Table 1 summarizes the key formulas for the variance ratios of forecasting errors between these two cases. We first note that if there is no secondary market effect, i.e., $\rho_{12} = \rho_{21} = 0$, then all the ratios in Table 1 are equal to one and therefore there is no difference in the forecasting accuracies between FIF and PIF. This is true because $\rho_{12} = \rho_{21} = 0$ implies $\theta_1 = \rho_{22}$ and $\theta_2 = \rho_{11}$, thus both the joint and the partial demand processes can be shown to reduce to the following same form:

$$z_{i,t} = \rho_{ii} z_{i,t-1} + \varepsilon_{i,t}, \quad \forall i = 1, 2.$$

(Note that the above equations are not equivalent to those associated with a pair of independent AR(1) processes because the error terms are contemporaneously correlated.⁵) Therefore, in the absence of a secondary market effect, FIF does not improve forecasting accuracy even if there is a strong primary market effect. In the rest of the section, we assume $\rho_{12} \neq 0$ and $\rho_{21} \neq 0$ and identify the conditions under which each party in a supply chain can benefit from FIF relative to PIF.

Theorem 4: *Given $\rho_{12} \neq 0$ and $\rho_{21} \neq 0$, FIF produces more accurate lead-time demand forecast for both retailers than PIF.*

Proof: See the appendix.

⁵ In the econometrics literature, this is referred to as a seemingly unrelated regression (SUR) model. See Greene (1993, p.674).

While FIF always benefits the retailers, it only improves the forecasting accuracy for the suppliers if an additional condition is imposed on the demand autocorrelations.

Theorem 5: *Given $\rho_{12} \neq 0$ and $\rho_{21} \neq 0$, FIF produces more accurate forecasts for each supplier if and only if $\rho_{11} + \rho_{22} > 0$.*

Proof: See the appendix.

Theorem 5 implies that more information, e.g., the external demand history of the other supply chain, does not automatically guarantee benefits for every party in a supply chain. FIF changes the suppliers' order processes and can make them less predictable, depending on the sign of the sum of demand autocorrelation coefficients. Our finding is related to Lee et al. (1997) that considers a single supply chain facing an AR(1) demand, and shows that the bullwhip effect exists when demand autocorrelation is positive. In their context, two forces drive the bullwhip effect. First, the order-up-to level policy is time-varying and the policy parameters chase demand through forecasting updates. Second, the MMSE forecasting procedure produces lead-time demand forecasts that are a linear combination of the observed demand. Both of these two forces remain in play in our setting. A third force is introduced here. Additional demand information alters the dynamic structure and the innovation error of the demand process identifiable by the retailer. When combined, these forces may destabilize the order process (the bullwhip effect) and make it harder for the supplier to forecast.

Another intuitive explanation for Theorem 5 can be gleaned from Miyaoka and Hausman (2004). They show that in a serial system, using more dated demand history for forecasting can stabilize the order process and create value for the supplier. Under FIF, each retailer only uses the most recent demand information from both retailers; under PIF, the partial demand forecasts are based on the partial demand process which use two previous data points plus the observed error. When $\rho_{11} + \rho_{22} > 0$, the partial demand forecasts can be viewed as being made with more dated demand information than FIF, leading to a more stable order process.

Theorem 6: *For each supply chain, the sum of the retailer's and supplier's forecasting error variances under FIF is not greater than that under PIF.*

Proof: See the appendix.

Theorem 6 implies that FIF always reduces the sum of the forecasting error variance for each supply chain. Together, Theorems 5 and 6 imply that FIF creates win-win situations when $\rho_{11} + \rho_{22} > 0$. When $\rho_{11} + \rho_{22} < 0$, however, FIF can cause conflicting economic consequences between the supplier and the retailer, even though the retailer's gain in forecasting accuracy outweighs the supplier's loss and the overall supply chain forecasting accuracy benefits from FIF. This may not present an obstacle for implementing FIF if the retailer's gain can be redistributed to compensate the supplier. Such a compensation mechanism can be designed more easily if both the retailer and the supplier in a serial chain belong to the same company. Otherwise, contracts may have to be negotiated to provide the incentive for the supplier to go along with FIF.

6.2. FIF vs. DIF

In this section, we focus on the special case of a one-period delay for DIP due to its analytical tractability. For longer delay periods, analytical solutions are much more difficult to obtain since the complex systems of nonlinear equations in Eq. (32) must be solved to determine the demand process under DIF. Fortunately, many of the insights can be gleaned from the special case, see numerical results in Section 6.4 on how some of the results here can be extrapolated beyond a one-period delay.

When the delay is one period long, the delayed demand process from Theorem 3 becomes

$$\begin{bmatrix} z_{1,t} \\ z_{2,t-1} \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12}\rho_{22} \\ \rho_{21} & \rho_{22} \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-2} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1,t} \\ \tilde{\varepsilon}_{2,t} \end{bmatrix} + \begin{bmatrix} \tilde{\theta}_{1,1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\varepsilon}_{1,t-1} \\ \tilde{\varepsilon}_{2,t-1} \end{bmatrix},$$

where $\tilde{\varepsilon}_{1,t} - \tilde{\theta}_{1,1}\tilde{\varepsilon}_{1,t-1} = \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1}$ and the moving average coefficients $\tilde{\theta}_{1,1}$ and $V(\tilde{\varepsilon}_{1,t}) = \tilde{\sigma}_{11}^2$ are obtained from solving the following nonlinear equations:

$$\begin{cases} (1 + \tilde{\theta}_{1,1}^2)\tilde{\sigma}_{11}^2 = V(\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1}) = \sigma_{11}^2 + \rho_{12}^2\sigma_{22}^2 \\ -\tilde{\theta}_{1,1}\tilde{\sigma}_{11}^2 = COV(\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1}, \varepsilon_{1,t-1} + \rho_{12}\varepsilon_{2,t-2}) = \rho_{12}\sigma_{12} \end{cases} \quad (36)$$

Clearly, the moving average part of the delayed demand process for $z_{1,t}$ has an identical structure to the PIF, except that we have replaced error $v_{1,t}$ with $\tilde{v}_{1,t} = \varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1}$. Table 2 summarizes

the forecasting errors and their variances corresponding to a one-period delay. We state Theorems 7-9 without proofs as they are similar to those of Theorems 4-6.

Theorem 7: *FIF produces more accurate lead-time demand forecasts for both retailers than DIF.*

Theorem 8: *FIF produces more accurate forecasts than DIF for supplier $i=1, 2$ if and only if $\rho_{ii} > 0$.*

The flip side of Theorem 8, namely, delayed demand information can be beneficial to the upstream suppliers, has been reported in other studies on a single supply chain. Zhang (2005) reports that in a serial supply chain, delayed AR(1) demand data for forecasting update can dampen the bullwhip effect. As mentioned before, Miyaoka and Hausman (2004) demonstrate that using forecasting results based on more dated demand data can be beneficial to the supply chain. The fundamental insight from these two studies is that when demand data is current, time-series forecasts tend to adjust temporary demand fluctuations in real time, which can destabilize the orders. On the other hand, stale demand data makes the demand forecasts less responsive to temporary demand changes, which can stabilize the orders. Our result suggests that this insight remains valid in a two demand streams setting where there is a lag in the availability of the other demand stream. The mechanism behind the stabilizing effect, however, is different, as we have pointed out in Section 6.1. A retailer's perceived demand process as well as the associated error are altered by delaying the availability of demand history of the other retailer.

Theorem 9: *For each supply chain, the sum of the retailer's and supplier's forecasting error variances under FIF is always smaller than that under DIF.*

6.3. PIF vs. DIF

We compare the special case of DIF with a one-period delay with PIF. Given that PIF is the limiting case of DIF when the number of periods delayed approaches infinity, the comparisons here

allow us to gauge the impact of delaying demand information beyond one period. The joint demand process under DIF immediately yields the following result.

Theorem 10: For retailer i , a one-period delay in the availability of the demand data of retailer j , $j \neq i$, affects all parties of the supply chain in the same way as any periods of delay if and only if $\rho_{ji} = 0$.

Proof: Omitted.

Theorem 11: For retailer $i=1, 2$, PIF yields a larger forecasting error variance than DIF with a one-period delay if and only if $V(\tilde{v}_{i,t}) > V(v_{i,t})$.

Proof: See the appendix.

Theorems 10-11 imply that additional delay beyond one-period under DIF may or may not cause damage to the retailers. It depends critically on the demand autocorrelation coefficient of the other retailer. When it is zero, the resulting increase in forecasting inaccuracy from delaying one period is the same as the increase from delaying any period. For Retailer 1, the necessary and sufficient condition $V(\tilde{v}_{1,t}) > V(v_{1,t})$ in Theorem 11 is equivalent to $\rho_{22}(\rho_{22}\sigma_{11}^2 - 2\rho_{12}\sigma_{12}) > 0$, namely, the demand autocorrelation coefficient of the other retailer must possess the same sign as $(\rho_{22}\sigma_{11}^2 - 2\rho_{12}\sigma_{12})$. In the most interesting case when $\rho_{12}\sigma_{12} > 0$, if ρ_{22} is greater than a positive threshold given by $2\rho_{12}\sigma_{12}/\sigma_{11}^2$, PIF generates less accurate forecast than DIF with a one-period delay; otherwise, if ρ_{22} is less than this positive threshold, PIF is less damaging than DIF with a one-period delay.

Theorem 12: For supplier $i=1, 2$, PIF yields a larger forecasting error variance than DIF if $\rho_{ii} < 0$ and $\rho_{11} + \rho_{22} > 0$.

Proof: See the appendix.

While Theorem 11 indicates that retailers do not always suffer from longer delays beyond one-period, Theorem 12 establishes a similar result for the suppliers. The required conditions

state that for supplier i , only if the autocorrelation of the other demand stream dominates its own negative autocorrelation, then PIF yields a less accurate forecast than DIF with a one period-delay. With an AR(1) demand process, it is well known that a negative demand autocorrelation tends to stabilize orders. Therefore, it makes intuitive sense that a strong positive demand autocorrelation of the other demand stream tends to destabilize retailer i 's orders, and when it dominates the stabilizing effect of retailer i 's external demand autocorrelation, delay information beyond one period makes the order less stable.

6.4. Numerical Results

In this section, we present a numerical study to compare FIF, PIF and DIF for various demand interaction, autocorrelation and error covariance σ_{12} . To reduce the dimensions of the demand parameters, we focus on symmetric cases where

$$\rho_{11} = \rho_{22} \quad \rho_{12} = \rho_{21} \quad \sigma_{11} = \sigma_{22}.$$

We first study the sensitivity of the forecast error variance ratios, for short, variance ratios, with respect to demand interaction, ρ_{12} , for two scenarios: one with a positive demand autocorrelation, and the other with a negative demand autocorrelation. Within each scenario, we then study the sensitivity of the variance ratios with respect to the error covariance σ_{12} . Two sets of variance ratios are examined: PIF versus FIF, as reported in the last column of Table 1, and DIF versus FIF, as reported in the last column of Table 2.

Figure 1 plots the variance ratios between PIF and FIF, for the retailers and the suppliers as a function of ρ_{12} at $\rho_{11}=0.5$, -0.5 and $\sigma_{11}=\sigma_{22}=10$. To evaluate the impact of the primary market effect, we consider two different co-variances of opposite signs with $\sigma_{12}=5$ or -5 . By Theorem 5, both suppliers benefit from FIF in this case as $\rho_{22}+\rho_{11} = 1 > 0$. To maintain covariance stationarity (cf. Section 3), ρ_{12} must be in $(-0.5, 0.5)$. Figure 1 reveals the following two insights when demand autocorrelation ρ_{11} is positive:

1. As the ρ_{12} deviates away from zero, which indicates a stronger demand interaction (or a stronger secondary market effect, equivalently), FIF becomes increasingly more beneficial for both the suppliers and the retailers. In the numerical examples, the retailers can achieve up to a 30% reduction in the forecast error variance while the suppliers can enjoy up to a 50% level.

2. The error covariance σ_{12} (the primary market effect) has an impact on the benefits from FIF. When $\rho_{12} > 0$, changing σ_{12} from 5 to -5 increases the gain in forecasting accuracy for both retailers and suppliers. When $\rho_{12} < 0$, however, changing σ_{12} from 5 to -5 decreases the gain in forecasting accuracy for both retailers and suppliers. Intuitively, FIF is likely to be more beneficial when there is a strong positive (negative) secondary market effect and a negative (positive) primary market effect.

In contrast when ρ_{11} is negative with a value of -0.5 , we observe that:

3. The retailers enjoy a similar range of benefits from FIF as the interaction coefficient deviates away from zero; however, the impact of changing σ_{12} from 5 to -5 is reversed, e.g., the curve corresponding to $\sigma_{12} = 5$ in the bottom panel of Figure 1 is identical to that of $\sigma_{12} = -5$ in the top panel of Figure 1.
4. For the suppliers, we observe that as the ρ_{12} deviates away from zero, FIF produces increasing loss in forecasting accuracy for either a negative or a positive σ_{12} . For example, when $\sigma_{12} = 5$, the loss can reach a level of nearly 80% as the interacting coefficient ρ_{12} approaches its lower limit of -0.5 . These observations indicate that FIF can have substantial negative impact on the suppliers when $\rho_{22} + \rho_{11} < 0$ and a strong market interaction is present.

Figure 2 examines the forecasting error variance ratio between DIF with a one-period delay and FIF. The layout is similar to Figure 1. The variance ratio associated with the retailer is a symmetric function of cross demand correlation σ_{12} , hence, we set it at a value of -3 rather than -5 to see how $|\sigma_{12}|$, i.e., the strength of cross correlation regardless of the direction of correlation, impacts the result. We observe that the retailer's forecasting accuracy increases in $|\rho_{12}|$ and decreases in $|\sigma_{12}|$ regardless of the sign of demand autocorrelation. The supplier's forecasting error variance ratios show a similar pattern as in Figure 1. By comparing the height of the corresponding curves in Figure 1 and Figure 2, we observe that delaying one period causes most of the incremental loss in forecasting accuracies for both retailers and suppliers, and further delays beyond one-period do not have as significant an impact.

7. Conclusions

In this paper, we address the question of when one supply chain's acquisition of the external demand information of another supply chain is beneficial. In a parallel structure of two serial sup-

ply chains with interacting demand streams, we develop the following insights through a combination of analytical and numerical studies:

1. When there is an interaction between the demand processes, the downstream parties of both supply chains can always reduce their forecast errors by availing themselves of the demand history of the other supply chain even with a delay; however, the same observation does not hold for the suppliers. Indeed, the conditions under which additional demand information improves the forecasting accuracy for the suppliers depend on demand interaction as well as on demand autocorrelation.
2. When full or delayed demand information is beneficial for all parties, our numerical results show that the forecast error variance reduction can be significant for both the retailers and the suppliers. This is especially true when there are strong interactions between the demand processes.
3. The overall forecasting accuracy of each supply chain is improved from acquiring the external demand information of the other supply chain.

From a managerial perspective, these findings have the following direct implications: First, the identifiable demand process for one retailer depends on what is known about the other interacting demand stream. Consequently, critical supply chain decisions that depend on demand forecasts are affected. Our results shed light on how the demand signal is altered by available information. Second, our results indicate that more demand information does not always automatically benefit all parties in a supply chain. In our setting, the overall gain for one supply chain is sometimes at the expense of the supplier. Our sufficient and necessary conditions on the win-win situations can help a manager to find synergy between interacting demand streams and offer guidelines on when it is appropriate to implement data acquisition programs.

These results, however, are based upon assumptions of zero transportation lead times and two parallel supply chains. Relaxing the first assumption would require retailers and suppliers to make forecasts beyond the immediate future period, and the forecasting errors at both stages of each supply chain are complex aggregates of future joint errors under FIF and future partial errors under PIF. Consequently, deriving explicit formulas for calculating the forecasting error variance remains a substantial challenge because the joint and partial errors are cross-correlated. Studying more-than-two parallel supply chains with interacting demand streams scales up our model exponentially because the growing number of demand information settings and the num-

ber of possible scenarios for demand parameters (i.e., combinations of interaction, across-correlation and auto-correlation coefficients) become combinatorial. Nevertheless, extension to non-zero lead times with multiple parallel supply chains represents a promising direction for future research.

Other supply system topology besides the parallel structure can be analyzed in our framework of multiple demand streams. In the manufacturing environment, a supplier may provide a common component to multiple competing manufacturers, while in a distribution system, a distributor may serve multiple competing retailers. In these cases, the distributor can aggregate information from multiple demand streams and share it with the retailers to gain additional forecasting benefits. Such a system warrants a study of its own because a distribution system is harder to analyze than serial supply chains, and the confidentiality issues must be addressed.

8. Appendix: Proofs

We only present proofs for Retailer 1, which are applicable to Retailer 2 with simple modifications.

Theorem 1

Retailer 1's partial demand process can be represented as an ARMA process if we can find a sequence of i.i.d errors that replicates the autocorrelation function of $v_{1,t} = \varepsilon_{1,t} - \rho_{22}\varepsilon_{1,t-1} + \rho_{12}\varepsilon_{2,t-1}$ at all lags. Based upon the joint error variance and covariance structure in Eq. (5), we have

$$\begin{aligned} V(v_{1,t}) &= \sigma_{11}^2 + \rho_{22}^2\sigma_{11}^2 + \rho_{12}^2\sigma_{22}^2 - 2\rho_{22}\rho_{12}\sigma_{12} \\ Cov(v_{1,t}, v_{1,t-1}) &= \rho_{12}\sigma_{12} - \rho_{22}\sigma_{11}^2, \end{aligned} \quad (37)$$

and $v_{1,t}$ is uncorrelated beyond first lag. Let $\xi_{1,t}$ be an i.i.d random error with a standard deviation of $\sigma_{1\xi}$. We formulate a first-order moving average error term $\xi_{1,t} - \theta_1\xi_{1,t-1}$. If we can find a $\sigma_{1\xi}$ and θ_1 that satisfy the following relationship:

$$\begin{cases} (1 + \theta_1^2)\sigma_{1\xi}^2 = V(v_{1,t}) \\ -\theta_1\sigma_{1\xi}^2 = COV(v_{1,t}, v_{1,t-1}), \end{cases} \quad (38)$$

the moving average error structure $\xi_{1,t} - \theta_1\xi_{1,t-1}$ replicates the entire autocorrelation function of $v_{1,t}$. This process then represents what Retailer 1 is capable of identifying using its own demand history. After dividing the second line in Eq. (38) into the first line to eliminate $\sigma_{1\xi}^2$, we obtain the following quadratic equation with θ_1 as the single unknown:⁶

⁶ Hamilton (1994, pp. 102-105) uses a similar quadratic equation to derive the time-series structure of the sum of two independent ARMA processes.

$$COV(v_{1,t}, v_{1,t-1})\theta_1^2 + V(v_{1,t})\theta_1 + COV(v_{1,t}, v_{1,t-1}) = 0. \quad (39)$$

The above quadratic equation yields two roots given by

$$\theta_1 = \frac{-V(v_{1,t}) \pm \sqrt{V^2(v_{1,t}) - 4COV^2(v_{1,t}, v_{1,t-1})}}{2COV(v_{1,t}, v_{1,t-1})}.$$

The term under the square root can be simplified to the following:

$$\begin{aligned} & V^2(v_{1,t}) - 4COV^2(v_{1,t}, v_{1,t-1}) \\ &= \left[\sigma_{11}^2 + \rho_{22}^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 - 2\rho_{22}\rho_{12}\sigma_{12} \right]^2 - 4(\rho_{12}\sigma_{12} - \rho_{22}\sigma_{11}^2)^2 \\ &= \left[(1 - \rho_{22})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1 - \rho_{22})\rho_{12}\sigma_{12} \right] \times \left[(1 + \rho_{22})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 - 2(1 + \rho_{22})\rho_{12}\sigma_{12} \right], \end{aligned}$$

which is always positive because the joint error variance-covariance matrix is positive definite. Therefore, both roots are real. The product of two roots is equal to one so they are reciprocal. Therefore, one solution is less than one in absolute value.

Let θ_l represent the real solution of Eq. (39) such that $|\theta_l| < 1$. Clearly, θ_l is uniquely determined by

$$\theta_l = \frac{-V(v_{1,t}) + \sqrt{V^2(v_{1,t}) - 4COV^2(v_{1,t}, v_{1,t-1})}}{2COV(v_{1,t}, v_{1,t-1})}. \quad (40)$$

The existence of this unique value allows us to represent the partial demand processes as invertible ARMA(2, 1) processes.

It is obvious that the partial process is covariance stationary because it shares the same characteristic equation as the joint demand process in Eq. (6). Solving for $\sigma_{1\xi}^2$ from the first line in Eq. (38) immediately produces the variance result stated in the theorem.

Theorem 2

The proof is greatly simplified with the lag operator B that shifts a time series variable backward in time. Here we have $Bz_{i,t} = Bz_{i,t-1}$ and $B\varepsilon_{i,t} = B\varepsilon_{i,t-1}$. We prove the result for Retailer 1 and the same argument remains valid for Retailer 2. As $l_1 \rightarrow \infty$, the term $\rho_{12}\rho_{22}^{l_1}z_{2,t-l_1-1} \rightarrow 0$ because the joint the process is assumed stationary and ρ_{22} is less than 1 in absolute value. Therefore, the delayed demand process in (29) approaches

$$\begin{aligned} z_{1,t} &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21} \left(z_{1,t-2} + \rho_{22}z_{1,t-3} + \dots + \rho_{22}^{l_1}z_{1,t-l_1} + \dots \right) \\ &\quad + \varepsilon_{1,t} + \rho_{12} \left(\varepsilon_{2,t-1} + \rho_{22}\varepsilon_{2,t-2} + \dots + \rho_{22}^{l_1}\varepsilon_{2,t-l_1-1} + \dots \right). \end{aligned}$$

The above can be represented in terms of B such that

$$\begin{aligned} z_{1,t} &= \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21} \left(1 + \rho_{22}B + \rho_{22}^2B^2 + \dots + \rho_{22}^{l_1}B^{l_1} + \dots \right) z_{1,t-2} \\ &\quad + \varepsilon_{1,t} + \rho_{12} \left(1 + \rho_{22}B + \rho_{22}^2B^2 + \dots + \rho_{22}^{l_1}B^{l_1} + \dots \right) \varepsilon_{2,t-1} \\ &= \rho_{11}z_{1,t-1} + \frac{\rho_{12}\rho_{21}}{(1 - \rho_{22}B)} z_{1,t-2} + \varepsilon_{1,t} + \frac{\rho_{12}}{(1 - \rho_{22}B)} \varepsilon_{2,t-1}. \end{aligned} \quad (41)$$

Multiplying $(1 - \rho_{22}B)$ through both sides of the above equation, we obtain

$$(1 - \rho_{22}B)z_{1,t} = (1 - \rho_{22}B)\rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}z_{1,t-2} + (1 - \rho_{22}B)\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t-1},$$

which yields the partial demand process in (21) after rearranging terms.

Theorem 3

For Retailer 1, its demand process under DIF in (29) can be joined with the delayed demand stream from Retailer 2 to form the following dynamic system of simultaneous equations:

$$\begin{cases} z_{1,t} = \rho_{11}z_{1,t-1} + \rho_{12}\rho_{21}(z_{1,t-2} + \rho_{22}z_{1,t-3} + \dots + \rho_{22}^{l_1-1}z_{1,t-l_1-1}) + \rho_{12}\rho_{22}^{l_1}z_{2,t-l_1-1} + \tilde{v}_{1,t} \\ z_{2,t-l_1} = \rho_{21}z_{1,t-l_1-1} + \rho_{22}z_{2,t-l_1-1} + \varepsilon_{2,t-l_1}. \end{cases} \quad (42)$$

The second line is from the joint demand lagged by l_1 periods. The autoregressive part of the above system, when written in matrix form, is the same as the AR component of our joint VARMA model in (31). The error part, however, is not in the form of the joint VARMA process because the error $\tilde{v}_{1,t}$ is autocorrelated at lags $k \leq l_1$ with autocovariance given by $\gamma_1(k)$. To transform the system into a VARMA model, we need to find two error terms $(\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t})$ independent over time that satisfy the following three conditions:

1. $V(\tilde{\varepsilon}_{2,t}) = \sigma_{22}^2$
2. $(\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t})$ are cross-correlated contemporaneously with $COV(\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t}) = \rho_{22}^{l_1-1}\sigma_{12}$
3. $V(\tilde{\varepsilon}_{1,t} - \tilde{\theta}_{1,1}\tilde{\varepsilon}_{1,t-1} - \tilde{\theta}_{1,2}\tilde{\varepsilon}_{1,t-2} - \dots - \tilde{\theta}_{1,l_1}\tilde{\varepsilon}_{1,t-l_1}) = V(\tilde{v}_{1,t})$ and the moving average term

$$\tilde{\varepsilon}_{1,t} - \tilde{\theta}_{1,1}\tilde{\varepsilon}_{1,t-1} - \tilde{\theta}_{1,2}\tilde{\varepsilon}_{1,t-2} - \dots - \tilde{\theta}_{1,l_1}\tilde{\varepsilon}_{1,t-l_1} \text{ must replicate all } \gamma_1(k) \text{ at lags } 0 < k \leq l_1$$

When these three conditions are met, we have the required error specifications of a VARMA model that is equivalent to the original joint system in (42) because vector moving average formed from $(\tilde{\varepsilon}_{1,t}, \tilde{\varepsilon}_{2,t})$ in the theorem produce the same variance and covariance between $\tilde{v}_{1,t}$ and $\varepsilon_{2,t-l_1}$ at all lags. It is well known in the time-series literature (see Box et al. 1993, Section 3.3.2, p. 70) that an MA(l_1) term formed with i.i.d errors is required to generate auto-covariances that cuts off at lag l_1 and the associated moving average coefficients and the variance of the error term must satisfy nonlinear system of equations in (32).

Theorem 4: $\sigma_{1\xi}^2 \geq \sigma_{11}^2$

Eliminating θ_1 from Eq. (38), we can obtain the following quadratic equation for solving $\sigma_{1\xi}^2$:

$$(\sigma_{1\xi}^2)^2 - V(v_{1,t})\sigma_{1\xi}^2 + [Cov(v_{1,t}, v_{1,t-1})]^2 = 0. \quad (43)$$

The sum of the two solutions must equal $V(v_{1,t})$. To be consistent with our choice of $|\theta_1| < 1$, we need to take the following as our solution for $\sigma_{1\xi}^2$, which is the larger of the two solutions:

$$\sigma_{1\xi}^2 = \frac{V(v_{1,t}) + \sqrt{[V(v_{1,t})]^2 - 4[Cov(v_{1,t}, v_{1,t-1})]^2}}{2}$$

Substituting Eq. (37) in the above equation, the term under the square root term can be expanded into the following:

$$\begin{aligned}
& \sqrt{[V(v_{1,t})]^2 - 4[Cov(v_{1,t}, v_{1,t-1})]^2} \\
&= \sqrt{(1-\rho_{22})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1-\rho_{22})\rho_{12}\sigma_{12}} \\
& \times \sqrt{(1+\rho_{22})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 - 2(1+\rho_{22})\rho_{12}\sigma_{12}} \\
&= \sqrt{E\left([\varepsilon_{1,t} - \rho_{12}\varepsilon_{2,t}]^2\right)} \times \sqrt{E\left([\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t}]^2\right)}.
\end{aligned}$$

Based on Schwarz's inequality (a special case of Holder's inequality), the last line in the above satisfies the following inequality:

$$\begin{aligned}
& \sqrt{E\left([\varepsilon_{1,t} - \rho_{12}\varepsilon_{2,t}]^2\right)} \times \sqrt{E\left([\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t}]^2\right)} \\
& \geq E\left([\varepsilon_{1,t} - \rho_{12}\varepsilon_{2,t}][\varepsilon_{1,t} + \rho_{12}\varepsilon_{2,t}]\right) \\
&= E\left[(1-\rho_{22})\varepsilon_{1,t}^2 - \rho_{12}^2\varepsilon_{2,t}^2 + 2\rho_{12}\rho_{22}\varepsilon_{1,t}\varepsilon_{2,t}\right] \\
&= (1-\rho_{22}^2)\sigma_{11}^2 - \rho_{12}^2\sigma_{22}^2 + 2\rho_{12}\rho_{22}\sigma_{12}.
\end{aligned}$$

The equality holds if and only if $\rho_{12} = 0$. Therefore, we have

$$\begin{aligned}
\sigma_{1\xi}^2 &= \frac{V(v_{1,t}) + \sqrt{[V(v_{1,t})]^2 - 4[Cov(v_{1,t}, v_{1,t-1})]^2}}{2} \\
&\geq \frac{V(v_{1,t}) + (1-\rho_{22}^2)\sigma_{11}^2 - \rho_{12}^2\sigma_{22}^2 + 2\rho_{12}\rho_{22}\sigma_{12}}{2} \\
&= \sigma_{11}^2,
\end{aligned}$$

with $\sigma_{1\xi}^2 = \sigma_{11}^2$ if only if $\rho_{12} = 0$.

Theorem 5

We need to show that $(1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2 > (1 + \rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12}$ if and only if $\rho_{11} + \rho_{22} > 0$ and $\rho_{12} \neq 0$. If we expand the squared term $(1 + \beta_1 - \theta_1)^2$, we have

$$(1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2 = (1 + \beta_1)^2 \sigma_{1\xi}^2 + \theta_1^2 \sigma_{1\xi}^2 - 2(1 + \beta_1)\theta_1 \sigma_{1\xi}^2,$$

Substituting from Eq. (38) and simplifying yield

$$\begin{aligned}
(1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2 &= (1 + \beta_1)^2 \sigma_{1\xi}^2 + V(v_{1,t}) - \sigma_{1\xi}^2 + 2(1 + \beta_1)(\rho_{12}\sigma_{12} - \rho_{22}\sigma_{11}^2) \\
&= \left\{[(1 + \beta_1)^2 - 1]\sigma_{1\xi}^2 + \sigma_{11}^2 - \rho_{22}^2\sigma_{11}^2 - 2(1 + \rho_{11})\rho_{22}\sigma_{11}^2\right\} + \rho_{12}^2\sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12}
\end{aligned} \tag{44}$$

,because $\sigma_{1\xi}^2 \geq \sigma_{11}^2$ from Theorem 4, the term in the bracelets on the last line satisfies the following inequality if and only if $\rho_{11} + \rho_{22} > 0$ and $\rho_{12} \neq 0$:

$$\begin{aligned}
& \left[(1 + \beta_1)^2 - 1\right]\sigma_{1\xi}^2 + \sigma_{11}^2 - \rho_{22}^2\sigma_{11}^2 - 2(1 + \rho_{11})\rho_{22}\sigma_{11}^2 \\
& > \left[(1 + \beta_1)^2 - 1\right]\sigma_{11}^2 + \sigma_{11}^2 - \rho_{22}^2\sigma_{11}^2 - 2(1 + \rho_{11})\rho_{22}\sigma_{11}^2 \\
& = (1 + \rho_{11})^2 \sigma_{11}^2.
\end{aligned}$$

Hence, the statement in Theorem 5 holds.

Theorem 6

The sum of the variances under FIF and PIF are

$$\text{Var. Sum FIF} = \sigma_{11}^2 + (1 + \rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12}$$

$$\text{Var. Sum PIF} = \sigma_{1\xi}^2 + (1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2,$$

Substituting $(1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2$ from (44) into the above and take the difference between the two variance sums, we have

$$\begin{aligned} & \text{Var. Sum PIF} - \text{Var. Sum FIF} \\ &= \left\{ \left[(1 + \beta_1)^2 - 1 \right] \sigma_{1\xi}^2 + \sigma_{11}^2 - \rho_{22}^2 \sigma_{11}^2 - 2(1 + \rho_{11})\rho_{22}\sigma_{11}^2 \right\} + \rho_{12}^2 \sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12} \\ &+ \sigma_{1\xi}^2 - \sigma_{11}^2 - \left[(1 + \rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1 + \rho_{11})\rho_{12}\sigma_{12} \right] \\ &= (1 + \beta_1)^2 (\sigma_{1\xi}^2 - \sigma_{11}^2), \end{aligned}$$

which is nonnegative and it is equal to zero when $\rho_{12} = 0$ (so $\sigma_{1\xi}^2 = \sigma_{11}^2$) or $\beta_1 = \rho_{11} + \rho_{22} = -1$.

Theorem 11

We denote $V = V(v_{1,t})$, $C = \text{Cov}(v_{1,t}, v_{1,t-1})$, $\tilde{V} = V(\tilde{v}_{1,t})$, and $\tilde{C} = \text{Cov}(\tilde{v}_{1,t}, \tilde{v}_{1,t-1})$. It is easy to verify the following identify from the definition of errors $v_{1,t}$ and $\tilde{v}_{1,t}$

$$C^2 - \tilde{C}^2 = (V - \tilde{V})\sigma_{11}^2. \quad (45)$$

Similar to the solution for $\sigma_{1\xi}^2$ under PIF, the variance $\tilde{\sigma}_{11}^2$ satisfy a quadratic equation similar to (43)

$$(\tilde{\sigma}_{11}^2)^2 - \tilde{V}\tilde{\sigma}_{11}^2 + \tilde{C}^2 = 0,$$

and we can obtain

$$\tilde{\sigma}_{11}^2 = \frac{\tilde{V} + \sqrt{\tilde{V}^2 - 4\tilde{C}^2}}{2}.$$

Subtracting the quadratic solutions for $\sigma_{1\xi}^2$ and $\tilde{\sigma}_{11}^2$ yields

$$\begin{aligned} 2(\sigma_{1\xi}^2 - \tilde{\sigma}_{11}^2) &= V - \tilde{V} + \sqrt{V^2 - 4C^2} - \sqrt{\tilde{V}^2 - 4\tilde{C}^2} \\ &= V - \tilde{V} + \frac{V^2 - \tilde{V}^2 - 4(C^2 - \tilde{C}^2)}{\sqrt{V^2 - 4C^2} + \sqrt{\tilde{V}^2 - 4\tilde{C}^2}}. \end{aligned}$$

Substituting (45) into the numerator of the above expression, we can simplify it into

$$2(\sigma_{1\xi}^2 - \tilde{\sigma}_{11}^2) = (V - \tilde{V}) \left[1 + \frac{V + \tilde{V} - 4\sigma_{11}^2}{\sqrt{V^2 - 4C^2} + \sqrt{\tilde{V}^2 - 4\tilde{C}^2}} \right]. \quad (46)$$

Substituting (45) into the numerator of the above expression and recognizing the following identities

$$\sqrt{V^2 - 4C^2} = 2\sigma_{1\xi}^2 - V, \quad \sqrt{\tilde{V}^2 - 4\tilde{C}^2} = 2\tilde{\sigma}_{11}^2 - \tilde{V}.$$

we can simplify (46) into

$$2(\sigma_{1\xi}^2 - \tilde{\sigma}_{11}^2) = 2(V - \tilde{V}) \left\{ \frac{(\sigma_{1\xi}^2 - \sigma_{11}^2) + (\tilde{\sigma}_{11}^2 - \sigma_{11}^2)}{\sqrt{V^2 - 4C^2} + \sqrt{\tilde{V}^2 - 4\tilde{C}^2}} \right\}.$$

The ratio term in the bracelets is always positive when there is interaction, hence, our conclusion stated in the theorem holds.

Theorem 12

The difference in Supplier 1's forecasting error variance is equal to

$$\begin{aligned} & (1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2 - (1 + \rho_{11} - \tilde{\theta}_{11})^2 \tilde{\sigma}_{11}^2 \\ &= \left[(1 + \theta_1^2) \sigma_{1\xi}^2 - 2(1 + \beta_1) \theta_1 \sigma_{1\xi}^2 + (\beta_1^2 + 2\beta_1) \sigma_{1\xi}^2 \right] \\ & \quad - \left[(1 + \tilde{\theta}_{11}^2) \tilde{\sigma}_{11}^2 - 2(1 + \rho_{11}) \tilde{\theta}_{11} \tilde{\sigma}_{11}^2 + (\rho_{11}^2 + 2\rho_{11}) \tilde{\sigma}_{11}^2 \right]. \end{aligned}$$

Substituting from (36) and (38) the terms $(1 + \tilde{\theta}_{11}^2) \tilde{\sigma}_{11}^2$, $\tilde{\theta}_{11} \tilde{\sigma}_{11}^2$, $(1 + \theta_1^2) \sigma_{1\xi}^2$, and $\theta_1 \sigma_{1\xi}^2$ into the above, we have

$$\begin{aligned} & (1 + \beta_1 - \theta_1)^2 \sigma_{1\xi}^2 - (1 + \rho_{11} - \tilde{\theta}_{11})^2 \tilde{\sigma}_{11}^2 \\ &= V(v_{1,t}) - V(\tilde{v}_{1,t}) + 2(1 + \rho_{11}) [COV(v_{1,t}, v_{1,t-1}) - COV(\tilde{v}_{1,t}, \tilde{v}_{1,t-1})] + 2\rho_{22} COV(v_{1,t}, v_{1,t-1}) \\ & \quad + (\beta_1^2 + 2\beta_1) \sigma_{1\xi}^2 - (\rho_{11}^2 + 2\rho_{11}) \tilde{\sigma}_{11}^2 \\ &= \rho_{22}^2 \sigma_{11}^2 - 2\rho_{22} \rho_{12} \sigma_{12} - 2(1 + \rho_{11}) \rho_{22} \sigma_{11}^2 + 2\rho_{22} (\rho_{12} \sigma_{12} - \rho_{22} \sigma_{11}^2) \\ & \quad + (\beta_1^2 + 2\beta_1) \sigma_{1\xi}^2 - (\rho_{11}^2 + 2\rho_{11}) \tilde{\sigma}_{11}^2. \end{aligned}$$

The last line of the above expressive can be simplified into

$$\left[(1 + \beta_1)^2 - 1 \right] (\sigma_{1\xi}^2 - \sigma_{11}^2) - \left[(1 + \rho_{11})^2 - 1 \right] (\tilde{\sigma}_{11}^2 - \sigma_{11}^2) \quad (47)$$

We know that $(\sigma_{1\xi}^2 - \sigma_{11}^2)$, and $(\tilde{\sigma}_{11}^2 - \sigma_{11}^2)$ are both positive when there is demand interaction, hence, both terms in (47) are positive if $\rho_{11} < 0$ and $\beta_1 = \rho_{11} + \rho_{22} > 0$.

Acknowledgement: The authors would like to thank the associate editor and two referees for their constructive comments and suggestions that have allowed us to significantly improve the paper. We also extend our gratitude to Rose Twomey for reviewing an early version of the paper and Carol Waller, a technical editor at Georgia Southern University, for editing and proof-reading during the review process. The second author is supported in part by the NSF contract CMMI-0747779.

References

- Aviv, Y. 2001. The effect of collaborative forecasting on supply chain performance. *Management Science* 47(10) 1326-1343.
- Aviv, Y. 2003. A time-series framework for supply-chain inventory management. *Operations Research* 51(2) 210-227.
- Aviv, Y., A. Federgruen, 1998. The operational benefits of information sharing and vendor managed inventory (VMI) programs. Working paper, Washington University in St. Louis, St. Louis, MO.
- Baganha, Manuel P. and Morris A. Cohen. 1998. The stabilizing effect of inventory in supply chains. *Operations Research* 46(3, supplement) S72-S83.
- Box, G. E. P., G. M. Jenkins, 1976. *Time Series Analysis: Forecasting and Control*, 2nd ed., Holden-Day, San Francisco.
- Bucklin, R. E., S. Gupta. 1999. Commercial use of UPC scanner data: Industry and academic perspectives. *Marketing Science* 18(3) 247-273.
- Cachon, G., M. Fisher. 2000. Supply chain inventory management and the value of shared information. *Management Science* 46(8) 1032-1048.
- Chen, Frank, Z. Drezner, J. K. Ryan, D. Simchi-Levi. 2000. Quantifying the bullwhip effect in a simple supply chain. *Management Science* 46(3) 436-443.
- Choi, H. P., J. D. Blocher, S. Gavirneni, 2008. Value of sharing production yield information in a serial supply chain. *Production and Operations Management* 17(6) 61-625.
- Ehrenberg, A. S. C. 1988. *Repeat-buying: Facts, Theory, and Applications*, Charles Griffin, London.
- Gaur, V., A. Giloni, S. Seshadri. 2005. Information sharing in a supply chain under ARMA demand. *Management Science* 51(6), 961-969.
- Gavirneni, S., R. Kapuscinski, S. Tayu. 1999. Value of information in capacitated supply chains. *Management Science* 45(1) 16-24.
- Gilbert, K. 2005. An ARIMA supply chain model. *Management Science* 51(2) 305-310.
- Graves, S. C. 1999. A single-item inventory model for a non-stationary demand process. *Manufacturing & Services Operations Management* 1(1) 50-61.
- Greene, W. H. 1993. *Econometric Analysis*, 3rd edition, Prentice Hall, Upper Saddle River, NJ.
- Hamilton, J. D. 1994. *Time Series Analysis*, Princeton University Press, Princeton, NJ.
- Lee, H. L., P. Padmanabhan, S. Whang. 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* 43(4) 546-58.
- Lee, H. L., K. C. So, C. S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* 46(5) 626-43.
- Lilien, G. L., P. Kotler, D. S. Moorthy. 1992. *Marketing Models*, Prentice Hall, Englewood Cliff, New Jersey.
- McFarlane, D., Y. Sheffi, 2008. The impact of automatic identification on supply chain operations. *The International Journal of Logistics Management* 14(1) 1-17.

- Mishra, B. K., S. Raghunathan, X. Yue. 2009. Demand forecast sharing in supply chains. *Production and Operations Management* 18(2) 152-166.
- Miyaoka, J., W. Hausman. 2004. How a base stock policy using “stale” forecasts provides supply chain benefits. *Manufacturing & Services Operations Management* 6(2) 149-162.
- Moinzadeh, K. 2002. A multi-echelon inventory system with information exchange. *Management Science* 48(3) 414-426.
- Morrison, D. G., D. C. Schmittlein. 1988. Generalizing the NBD model for customer purchases: What are the implications and is it worth the effort? *Journal of Business and Economics Statistics* 6(2) 145-159.
- Raghunathan, S. 2001. Information sharing in a supply chain: A note on its value when demand is non-stationary. *Management Science* 47(4) 605-610.
- Schultz, R. L., D. R. Wittink. 1976. The measurement of industry advertising effects. *Journal of Marketing Research* 13(1) 71-75.
- Simchi-Levi, D., Y. Zhao. 2003a. The impact of information sharing in a two-stage supply chain with production capacity constraints. *Naval Research Logistics* 50(8) 888-916.
- Simchi-Levi, D., Y. Zhao. 2003b. The impact of information sharing on forecasting accuracy in a multi-stage distribution system with stationary demand. Working Paper, Rutgers University, Newark, NJ.
- Varian, H. R. 1992. *Microeconomic Analysis*, W. W. Norton & Company, New York.
- van Ryzin, G., S. Mahajan. 1999. On the relationship between inventory costs and variety benefits in retail assortments. *Management Science* 45(11) 1496-1509.
- Wei, W. W. S. 1990. *Time Series Analysis: Univariate and Multivariate Methods*, Addison-Wesley.
- Zhang, X. 2004. Evolution of ARMA demand in supply chains, *Manufacturing and Services Operations Management* 6(2) 195–198.
- Zhang, X. 2005. Delayed demand information and dampened bullwhip effect. *Operations Research Letters* 33(3) 289-294.
- Zipkin, P. H. 2000. *Foundations of Inventory Management*, McGraw-Hill, New York.

Table 1: LT Demand Forecasting Errors and Their Variance Ratios under FIF and PIF

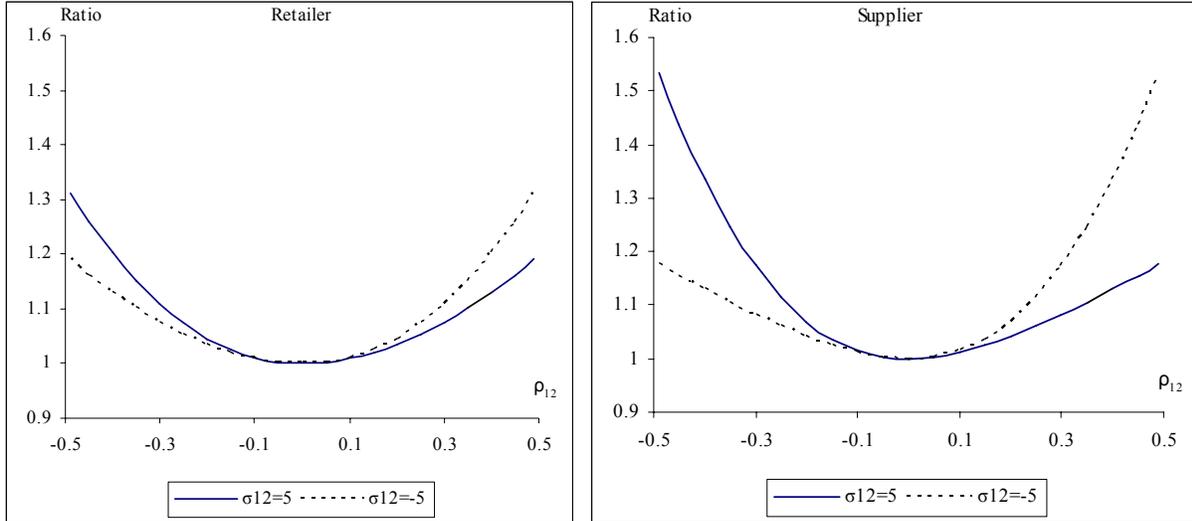
	<i>FIF</i>	<i>PIF</i>	<i>Ratio: PIF/FIF</i>
<i>Retailer 1</i>	$\varepsilon_{1,t+1}$	$\xi_{1,t+1}$	$\frac{\sigma_{1\xi}^2}{\sigma_{11}^2} = \frac{\sigma_{11}^2 + \rho_{22}^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 - 2\rho_{22}\rho_{12}\sigma_{12}}{(1+\theta_1^2)\sigma_{11}^2}$
<i>Retailer 2</i>	$\varepsilon_{2,t+1}$	$\xi_{2,t+1}$	$\frac{\sigma_{2\xi}^2}{\sigma_{22}^2} = \frac{\sigma_{22}^2 + \rho_{11}^2 \sigma_{22}^2 + \rho_{21}^2 \sigma_{11}^2 - 2\rho_{11}\rho_{21}\sigma_{12}}{(1+\theta_2^2)\sigma_{22}^2}$
<i>Supplier 1</i>	$(1+\rho_{11})\varepsilon_{1,t+1}$ $+\rho_{12}\varepsilon_{2,t+1}$	$(1+\beta_1-\theta_1)\xi_{1,t+1}$	$\frac{(1+\beta_1-\theta_1)^2 \sigma_{1\xi}^2}{(1+\rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1+\rho_{11})\rho_{12}\sigma_{12}}$
<i>Supplier 2</i>	$(1+\rho_{22})\varepsilon_{2,t+1}$ $+\rho_{21}\varepsilon_{1,t+1}$	$(1+\beta_1-\theta_2)\xi_{2,t+1}$	$\frac{(1+\beta_1-\theta_2)^2 \sigma_{2\xi}^2}{(1+\rho_{22})^2 \sigma_{22}^2 + \rho_{21}^2 \sigma_{11}^2 + 2(1+\rho_{22})\rho_{21}\sigma_{12}}$

Table 2: LT Demand Forecasting Errors and Their Variance Ratios under FIF and DIF

	<i>FIF</i>	<i>DIF</i>	<i>Ratio: DIF/FIF</i>
<i>Retailer 1</i>	$\varepsilon_{1,t+1}$	$\tilde{\xi}_{1,t+1}$	$\frac{\tilde{\sigma}_{11}^2}{\sigma_{11}^2} = \frac{\sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2}{(1+\tilde{\theta}_1^2)\sigma_{11}^2}$
<i>Retailer 2</i>	$\varepsilon_{2,t+1}$	$\tilde{\xi}_{2,t+1}$	$\frac{\tilde{\sigma}_{22}^2}{\sigma_{22}^2} = \frac{\sigma_{22}^2 + \rho_{21}^2 \sigma_{11}^2}{(1+\tilde{\theta}_2^2)\sigma_{22}^2}$
<i>Supplier 1</i>	$(1+\rho_{11})\varepsilon_{1,t+1}$ $+\rho_{12}\varepsilon_{2,t+1}$	$(1+\rho_{11}-\tilde{\theta}_{1,1})\tilde{\xi}_{1,t+1}$	$\frac{(1+\rho_{11}-\tilde{\theta}_{1,1})^2 \tilde{\sigma}_{11}^2}{(1+\rho_{11})^2 \sigma_{11}^2 + \rho_{12}^2 \sigma_{22}^2 + 2(1+\rho_{11})\rho_{12}\sigma_{12}}$
<i>Supplier 2</i>	$(1+\rho_{22})\varepsilon_{2,t+1}$ $+\rho_{21}\varepsilon_{1,t+1}$	$(1+\rho_{22}-\tilde{\theta}_{2,1})\tilde{\xi}_{2,t+1}$	$\frac{(1+\rho_{22}-\tilde{\theta}_{2,1})^2 \tilde{\sigma}_{22}^2}{(1+\rho_{22})^2 \sigma_{22}^2 + \rho_{21}^2 \sigma_{11}^2 + 2(1+\rho_{22})\rho_{21}\sigma_{12}}$

Figure 1: Forecasting Error Variance Ratios as a Function of Demand Interaction ρ_{12} –PIF versus FIF

Positive Demand Autocorrelation $\rho_{11}=0.5$



Negative Demand Autocorrelation $\rho_{11} = -0.5$

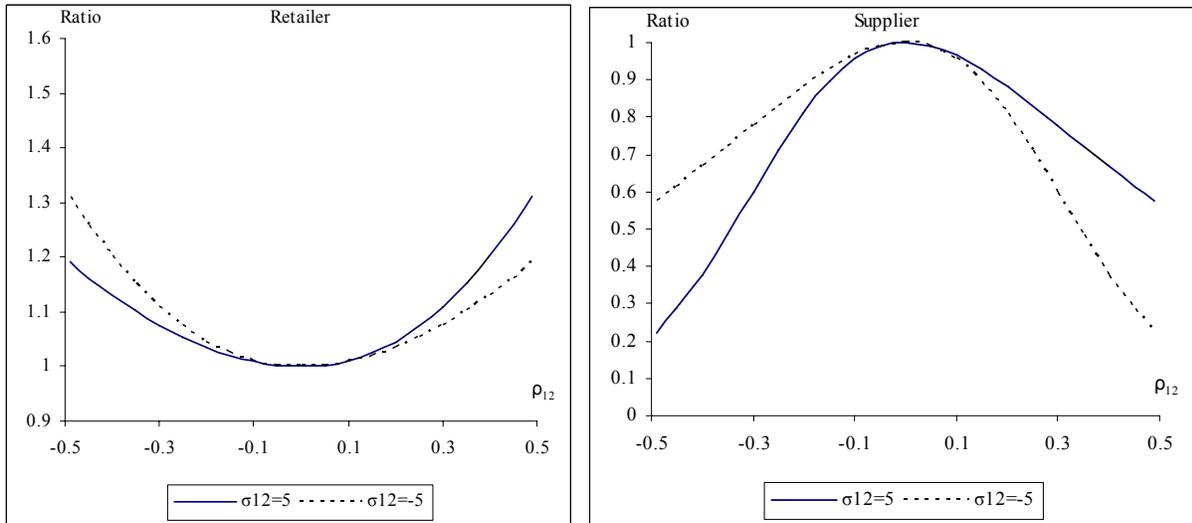
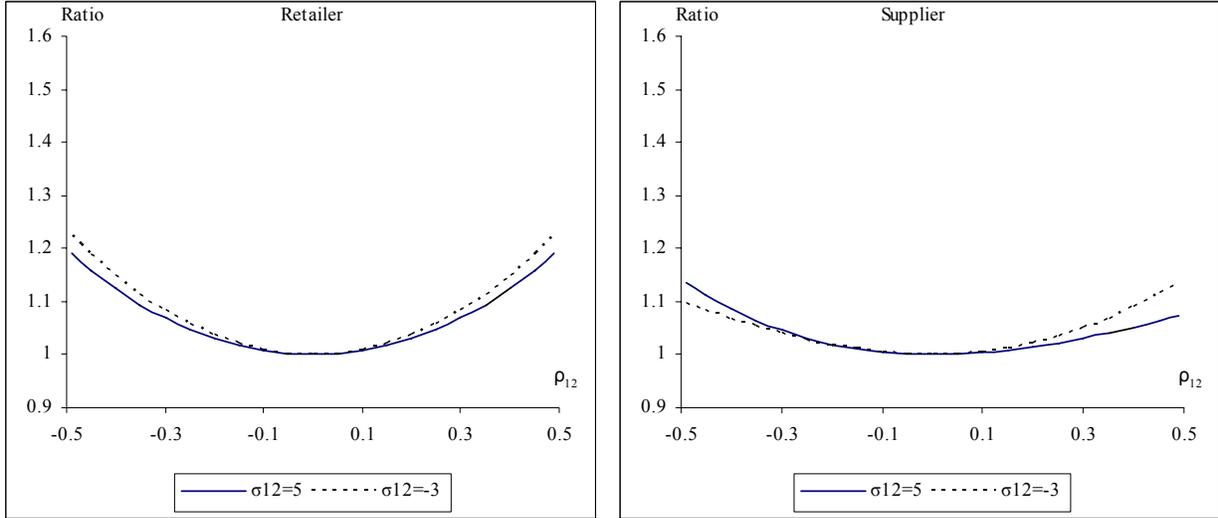


Figure 2: Forecasting Error Variance Ratios as a Function of Demand Interaction ρ_{12} –DIF versus FIF

Positive Demand Autocorrelation $\rho_{11}=0.5$



Negative Demand Autocorrelation $\rho_{11}=-0.5$

