Impact of Exponential Smoothing on Inventory Costs in Supply Chains

Meng-Chen Hsieh *, Avi Giloni†, Clifford Hurvich ‡

February 1, 2016

Abstract

It is common for firms to forecast stationary demand using simple exponential smoothing due to the ease of computation and understanding of the methodology. In this paper we show that the use of this methodology can be extremely costly in the context of inventory in a two-stage supply chain when the retailer faces AR(1) demand. We show that under the myopic order-up-to level policy, a retailer using exponential smoothing may have expected inventory-related costs more than ten times higher than when compared to using the optimal forecast. We demonstrate that when the AR(1) coefficient is less than the exponential smoothing parameter, the supplier's expected inventory-related cost is less when the retailer uses optimal forecasting as opposed to exponential smoothing. We show there exists an additional set of cases where the sum of the expected inventory-related costs of the retailer and the supplier is less when the retailer uses optimal forecasting as opposed to exponential smoothing even though the supplier's expected

*College of Business Administration, Department of Information Systems and Supply Chain Management, Rider University, Sweigart Hall 313, 2083 Lawrenceville Road, Lawrenceville, NJ 08648. E-mail: *mehsieh@rider.edu*

†Sy Syms School of Business, Yeshiva University, BH-428, 500 West 185th St., New York, NY 10033. E-mail: *agiloni@yu.edu*

‡Department of Information, Operations, and Management Science, Leonard N. Stern School of Business, New York University, Suite 8-52, 44 West 4th St., New York, NY 10012. E-mail: *churvich@stern.nyu.edu*

cost is higher. In this paper, we study the impact on the naive retailer, the sophisticated supplier, and the two-stage chain as a whole of the supplier sharing its forecasting expertise with the retailer. We provide explicit formulas for the supplier's demand and the mean squared forecast errors for both players under various scenarios.

1 Introduction

It is common for firms to forecast stationary demand using simple exponential smoothing for inventory control (see for example Nahmias 2015, p. 853) due to the ease of computation and understanding of the methodology. In this paper, we examine the impact of using this convenient forecasting procedure on inventory in supply chains. To do so, we consider a two-stage supply chain with a naive retailer that faces AR(1) demand but uses exponential smoothing for forecasting even though this yields suboptimal forecasts. We study the impact on the retailer, the supplier (assumed to be more sophisticated), and the chain as a whole of the supplier sharing its forecasting expertise with the retailer.

There has been much research on the value of information sharing in supply chains when the retailer faces AR (autoregressive) or ARMA (autoregressive moving average) demand. In their seminal paper, Lee, So, and Tang (2000) studied the value of information sharing in a two-stage supply chain where both players use the myopic order-up-to level policy and the retailer faces an AR(1) demand with positive AR coefficient. They concluded that there is always value to the supplier of the retailer sharing its demand. Raghunathan (2001) showed that in this case, the supplier is always able to infer the retailer's demand and hence there is no value to information sharing. Zhang (2004), Gaur, Giloni, and Seshadri (2005) hereafter referred to as GGS, and Giloni, Hurvich, and Seshadri (2014) (GHS hereafter) studied the more general ARMA demand case which included AR(1) demand with a negative AR(1) coefficient. Zhang assumed that the shock sequence in the retailer's order to the supplier would be observable to the supplier and hence concluded that there is no value to information sharing. GGS showed that when the AR(1) coefficient is less than $-.5$, the supplier is unable to recover the retailer's demand. GHS determined when there is value to information sharing (for example, the case discussed above), assuming that all players use optimal forecasts. In this framework, they also studied the propagation of demand (ARMA-in ARMA-out) up the supply chain with and without sharing of shocks. In this stream of research, the only information sharing considered is the retailer sharing demand information with the supplier. Others have studied environments where the retailer and/or the supplier can share demand information with each other. For example, Shnaiderman and El Ouardighi (2014) assumed that the retailer observes AR(1) demand where the random component of the retailer's demand is a function of both the retailer's and the supplier's information. They studied when it is beneficial for either player to share information with the other and when information sharing might be detrimental to a player.

In this paper, we consider the value to the retailer and the supplier of the supplier sharing its forecasting expertise in a two-stage supply chain. We assume that the retailer, for the sake of convenience, uses the widely available simple exponential smoothing method to forecast its demand. This creates a potential disconnect between the true mechanism generating demand (which we assume to be AR(1)) and the forecasting methodology. Indeed, unless the retailer's demand was generated by an ARIMA(0,1,1) model, the exponential smoothing forecast will be suboptimal. On the other hand, we assume that the supplier is sufficiently sophisticated in modeling and data analysis that the supplier, given a sufficiently long history of the retailer's orders, is able to infer the true ARMA model generating the retailer's order process. We prove (see Remark 1) that the supplier is then able to infer the retailer's demand as well as its AR(1) generating mechanism, and also the exponential smoothing parameter used by the retailer. Therefore, the supplier is in possession of expertise that would benefit the retailer.

Indeed, the retailer is always benefited by the use of the optimal forecast. However, as a result of the supplier sharing its forecasting expertise, we show that the demand the supplier will face can have a smaller or larger mean squared forecast error (MSFE) than when the retailer uses the suboptimal exponential smoothing forecast. We assume that the supplier will share its forecasting expertise when it benefits the supplier or when it benefits the chain as a whole (where the benefits will be shared equally between both players).

If the supplier provides its forecasting expertise to the retailer, there may be value in the retailer sharing its demand with the supplier. Specifically, we show that once the supplier shares its expertise with the retailer, who therefore now uses optimal forecasts, the supplier may no longer be able to recover the retailer's demand. In such a case, the supplier will benefit from the retailer sharing its demand with the supplier.

The remainder of this paper is organized as follows. In Section 2, we describe the setting of our two-stage supply chain in detail. In Section 3, we carry out a simulation study to gauge the inventory costs of the retailer, the supplier, and the whole supply chain. We consider different parameter configurations in the simulation and show cases where the supplier is better off by sharing its expertise with the retailer in terms of a lower inventory cost for itself or the whole chain and that the retailer always benefits if it uses the optimal forecast. In Section 4, we derive the retailer's order process based on its suboptimal (exponential smoothing) and optimal (AR(1)) forecast as well as the MSFE of these forecasts. We assume that the supplier has sufficient expertise to identify the form, model degree and coefficients of the retailer's order process correctly. We derive the MSFE of the supplier's best linear forecast under the two different retailer order processes. We present concluding remarks in Section 5.

2 Problem Setup

We consider a two-stage supply chain where there is one retailer and one supplier. Both players use the myopic order-up-to level policy to determine their inventory positions and hence their order quantities. We assume that the retailer's leadtime is ℓ_1 periods and that the supplier's leadtime is ℓ_2 periods. We assume that both players have a holding cost per unit per unit time of h and shortage cost per unit per unit time of s. At time t, the retailer observes its demand and then places an order with its supplier according to the myopic order-up-to level policy. In other words, the retailer's order-up-to-level at time t is

$$S_{1,t} = m_{1,t} + c\sqrt{\nu_1} \tag{1}$$

where ν_1 is the MSFE of $m_{1,t}$, the retailer's forecast of demand over the leadtime $t+1$ to $t+1+\ell_1$ based on its past and present demand at time t

$$m_{1,t} = \sum_{i=1}^{l_1+1} \hat{D}_{1,t+i}, \tag{2}$$

$\hat{D}_{1,t+i}$ is the retailer's forecast of demand at time $t+i$ based on $D_{1,t}, D_{1,t-1}, \ldots$ and $c = \Phi^{-1}(\frac{s}{s+h})$ (where $\Phi(\cdot)$ is the standard normal cumulative distribution function) is the fractile based upon service level $\frac{s}{s+h}$ and the distribution of the retailer's demand shocks.

The retailer's order to the supplier, $D_{2,t}$, is therefore equal to the current observed demand plus the order-up-to level at time t minus the order up-to level at time $t-1$ (see Lee, So, and Tang 2000, Equation (3.1)),

$$D_{2,t} = D_{1,t} + S_{1,t} - S_{1,t-1} = D_{1,t} + m_{1,t} - m_{1,t-1} \tag{3}$$

where the right hand side holds since the mean squared forecast error is time invariant. We refer the supplier's forecast of its demand over the leadtime $t+1$ to $t+1+\ell_2$ as $m_{2,t}$.

In this paper, we assume that the retailer's demand $\{D_{1,t}\}$ follows a stationary $AR(1)$ process with the representation

$$D_{1,t} = d + \phi D_{1,t-1} + \epsilon_{1,t}^{true} \tag{4}$$

where d is a constant, $|\phi| < 1$, and $\{\epsilon_{1,t}^{true}\}$ are the retailer's true shocks, which are Gaussian white noise with mean zero and variance $\sigma_{\epsilon_1}^2$. We are interested in studying how the value of ϕ in the retailer's demand model impacts its forecast of demand under exponential smoothing and hence how demand propagates upstream. To do so, we consider two forecasting methods used by the retailer. In the first case, the retailer uses exponential smoothing ($\hat{D}_{t+i}^{ES}$ given by Equations (16) and (35)). In the second case, upon the supplier providing forecasting expertise, the retailer uses the optimal forecast for its AR(1) demand ($\hat{D}_{t+i}^{AR}$ given by Equation (45)). We use Equation (3) to obtain the retailer's order to the supplier when the retailer uses exponential smoothing. We then determine the expected cost for the retailer when using exponential smoothing and the expected cost for the supplier whose demand is based upon the retailer's use of exponential smoothing. Finally, we compare the above to the case when the retailer uses the optimal forecast.

When the retailer uses exponential smoothing, we refer to its forecast over the leadtime as $m_{1,t}^{ES}$, its mean squared forecast error as ν_1^{ES}, and forecast for demand at time $t+i$ as $\hat{D}_{1,t+i}^{ES}$. When optimal forecasting is used, we replace the superscript ES by AR. Even though we provide the theoretical mean squared forecast error for a retailer using exponential smoothing (see Proposition 7), such a naive retailer would presumably not be able to derive this quantity. Instead, we assume that such a retailer would estimate this value by using the sample variance of its forecast errors over the leadtime, which converges to the true MSFE.

We define the following terms used in this paper:

• $m_{2,t}^{ES}$: supplier's best linear forecast of demand over the leadtime when it faces demand $\left\{D_{2,t}^{ES}\right\}$, i.e. the retailer uses exponential smoothing to forecast its demand.

• $m_{2,t}^{AR,S}$: supplier's best linear forecast of demand over the leadtime when it faces demand $\{D_{2,t}^{AR}\}$, i.e. the retailer uses the optimal forecast to predict its demand, and the supplier is able to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$ or the retailer shares its demand shocks with the supplier.

• $m_{2,t}^{AR,NS}$: supplier's best linear forecast of demand over the leadtime when it faces demand $\{D_{2,t}^{AR}\}$ and is not able to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$, and the retailer does not share its demand with the supplier.

• ν_2^{ES}: MSFE of the supplier's best linear forecast when the retailer uses exponential smoothing to forecast its demand.

• $\nu_2^{AR,S}$: MSFE of the supplier's best linear forecast when the retailer uses the optimal forecast to predict its demand and the supplier is able to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$, or the retailer shares its demand shocks with the supplier.

• $\nu_2^{AR,NS}$: MSFE of the supplier's best linear forecast when the retailer uses the optimal forecast to predict its demand and the supplier is unable to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$.

3 Inventory and Cost Implications of Exponential Smoothing

Both the retailer and the supplier want to minimize their own inventory-related costs. Let $S_{1,t}^{ES}$ denote the retailer's order-up-to level at time t if it adopts the suboptimal forecast and $S_{1,t}^{AR}$ the retailer's order-up-to at time t if it adopts the optimal forecast At time t, the retailer's actual inventory cost in period $t+\ell_1+1$ is given by

$$\mathrm{IC}_{1,t}^{ES} = \left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}^{ES}\right)^+ s + \left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}^{ES}\right)^- h \tag{5}$$

if the retailer uses exponential smoothing for its forecast and

$$\mathrm{IC}_{1,t}^{AR} = \left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}^{AR}\right)^+ s + \left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}^{AR}\right)^- h \tag{6}$$

if the retailer uses the optimal forecast. Following Lee, So, and Tang (2000) (see their Equation (4.7) and the surrounding discussion), we assume that the retailer, at time t, wishes to determine the value of $S_{1,t}$ to minimize its conditional expected inventory cost

$$E\left\{\left[\left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}\right)^+ s + \left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - S_{1,t}\right)^- h\right] \middle| \mathcal{M}_t^1\right\} \tag{7}$$

where $\mathcal{M}_t^1$ is retailer's available information at time t. Lee, So and Tang (2000) justified the order-up-to level (minimizing Equation (7))

$$S_{1,t}^* = m_{1,t}^* + \Phi^{-1}(\frac{s}{s+h})\sqrt{\nu_1^*} \tag{8}$$

where $m_{1,t}^*$ is the best linear forecast of leadtime demand and ν_1^* is the MSFE of $m_{1,t}^*$. Since we assume that the retailer faces AR(1) demand, $m_{1,t}^* = m_{1,t}^{AR}$, and $\nu_1^* = \nu_1^{AR}$.

As in Lee, So, and Tang (2000) (LST hereafter; see their Equation (4.8) and the discussion that precedes it), we assume that when the retailer uses a suboptimal forecast $m_{1,t}$ with MSFE ν_1, it will replace $m_{1,t}^*$ with $m_{1,t}$ and ν_1^* with ν_1 in Equation (8). In the context of this paper, the suboptimal forecast is $m_{1,t}^{ES}$ with corresponding MSFE ν_1^{ES}. In this section, we study the effect of the retailer's use of exponential smoothing on its inventory cost. We also consider the impact on the supplier of the retailer using exponential smoothing.

Consider the loss function

$$L(x) = \int_x^\infty (z-x)d\Phi(z) \tag{9}$$

where $\Phi(z)$ is the standard normal CDF. The retailer's optimal conditional expected cost in Proposition 1 below can be found in Equation (4.7) of LST. We include the proposition and its proof (in the Appendix) in our paper since they both are important for the other results in this section.

Proposition 1 *The retailer's optimal conditional expected cost is*

$$\sqrt{\nu_1^*}\left[(s+h)L\left(\Phi^{-1}\left(\frac{s}{s+h}\right)\right) + h\Phi^{-1}\left(\frac{s}{s+h}\right)\right]. \tag{10}$$

In the case where the retailer uses exponential smoothing, before we consider the conditional expected cost, we first discuss the conditional service level defined as $P\left(\sum_{i=1}^{\ell_1+1} D_{t+i} < S_{1,t}^{ES}|\mathcal{M}_t^1\right)$.

Proposition 2 *If the retailer uses the myopic order-up-to policy based on the exponential smoothing forecast, the retailer's conditional service level is given by $\Phi(r)$ where*

$$r = \frac{m_{1,t}^{ES} - m_{1,t}^* + \Phi^{-1}(\frac{s}{s+h})\sqrt{\nu_1^{ES}}}{\sqrt{\nu_1^*}}. \tag{11}$$

Proof : The retailer's conditional service level under exponential smoothing is given by (where Z is standard normal conditionally on $\mathcal{M}_t^1$)

$$\begin{aligned} P\left(\sum_{i=1}^{\ell_1+1} D_{t+i} < S_{1,t}^{ES}|\mathcal{M}_t^1\right) &= P\left(\frac{\sum_{i=1}^{\ell_1+1} D_{t+i} - m_{1,t}^*}{\sqrt{\nu_1^*}} < \frac{S_{1,t}^{ES} - m_{1,t}^*}{\sqrt{\nu_1^*}}\middle|\mathcal{M}_t^1\right) \\ &= P\left(Z < \frac{m_{1,t}^{ES} + \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}} - m_{1,t}^*}{\sqrt{\nu_1^*}}\middle|\mathcal{M}_t^1\right) \\ &= P\left(Z < r|\mathcal{M}_t^1\right) = \Phi(r). \end{aligned} \tag{12}$$

□

Note that when the retailer uses the optimal forecast, the service level is equal to $\frac{s}{s+h}$. Therefore, the conditional service level is always equal to $\frac{s}{s+h}$. However, Proposition 2 implies that when the retailer uses the exponential smoothing forecast, the retailer's conditional service level is the random quantity $\Phi(r)$, which in general will not be equal to the desired $\frac{s}{s+h}$.

The retailer's conditional expected cost under a suboptimal forecast is also discussed by LST in their Equation (4.9). We provide the retailer's conditional expected cost under exponential smoothing in the next proposition.

Proposition 3 *If the retailer uses the myopic order-up-to policy with the exponential smoothing forecast, the retailer's conditional expected cost is*

$$\sqrt{\nu_1^*}[(s+h)L(r) + hr]. \tag{13}$$

To demonstrate the effect of exponential smoothing on the retailer's conditional service level and its conditional expected cost, we simulated 1000 different paths of $AR(1)$ processes up to time $t = 100$. For each of the 1000 paths, we simulated 10000 possible values of $D_{1,101}$. We then computed the proportion of the 10000 realizations where the retailer would satisfy its leadtime demand using a myopic order-up-to-policy with $h = 1$ and $s = 9$ based upon an exponential smoothing forecast. We show the distribution of the retailer's conditional service level, $\Phi(r)$, in Figure 1. It can be seen that the retailer's conditional service level is random and there are a non-negligible number of observations where $\Phi(r)$ is less than the desired service level 0.9. In addition, for each of the 1000 paths, we computed the retailer's average inventory cost across the 10000 realizations and graphed it against the retailer's conditional service level. The bottom plot of Figure 1 shows a U-shaped curve of the retailer's conditional expected inventory cost versus its conditional service level. The minimum inventory cost occurs at service level 0.9. These graphs demonstrate that the exponential smoothing forecast when used within a myopic order-up-to policy will often cause the retailer to either overshoot or undershoot its optimal service level and thus drive up its conditional expected inventory cost.

Although the myopic-order-up-to policy is focused on minimizing the conditional expected cost, a manager is likely to measure the effectiveness of his inventory policy by considering the long-run average inventory cost per period. We analyze this (unconditional) expected cost for the retailer under optimal forecasting as well as under exponential smoothing. We define $E\text{IC}_1^{ES} = E\left[IC_{1,t}^{ES}\right]$ and $E\text{IC}_1^{AR} = E\left[IC_{1,t}^{AR}\right]$. Since by Proposition 1 the conditional expected cost under the optimal policy is not random, $E\text{IC}_1^{AR} = \sqrt{\nu_1^*}[(s+h)L(\Phi^{-1}(\frac{s}{s+h})) + h\Phi^{-1}(\frac{s}{s+h})]$. We next present the retailer's expected cost under exponential smoothing. As far as we are aware this result has not been explicitly presented in previous literature.

Proposition 4 *If the retailer uses the myopic order-up-to policy based on the exponential smooth-*

ing forecast, the retailer's expected cost is

$$EIC_1^{ES} = \sqrt{\nu_1^{ES}} \left[(s+h)L\left(\Phi^{-1}\left(\frac{s}{s+h}\right)\right) + h\Phi^{-1}\left(\frac{s}{s+h}\right) \right]. \tag{14}$$

We next consider the supplier's problem. A proof very similar to that of Proposition 1 shows that the supplier's optimal conditional expected cost is given by

$$\sqrt{\nu_2^*} \left[(s+h)L\left(\Phi^{-1}\left(\frac{s}{s+h}\right)\right) + h\Phi^{-1}\left(\frac{s}{s+h}\right) \right], \tag{15}$$

where ν_2^* is defined below. As discussed previously, we assume that the supplier always uses optimal forecasting. Nevertheless, the demand it observes depends upon the forecasting method used by the retailer as well as the demand sharing arrangement between the retailer and supplier. The quantity ν_2^* in Equation (15) is given by ν_2^{ES} (when the retailer uses exponential smoothing to forecast its demand), or by $\nu_2^{AR,S}$ (when the retailer uses the optimal forecast to predict its demand and the supplier is able to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$, or the retailer shares its demand shocks with the supplier), or by $\nu_2^{AR,NS}$ (when the retailer uses the optimal forecast to predict its demand, the supplier is unable to infer the retailer's demand shocks from $\{D_{2,t}^{AR}\}$ and the retailer does not share its demand shocks with the supplier).

Similarly, the expected total cost of the supply chain when the retailer uses exponential smoothing is given by $TC^{ES} = E\mathrm{IC}_1^{ES} + E\mathrm{IC}_2^{ES}$. The expected total cost of the supply chain when the retailer uses optimal forecasting and the retailer shares its demand shocks or the supplier is able to infer them is given $TC^{CS} = E\mathrm{IC}_1^{AR} + E\mathrm{IC}_2^{AR,S}$. Finally, the expected total cost of the supply chain when the retailer uses optimal forecasting and the retailer does not share its demand shocks and the supplier is not able to infer the retailer's demand shocks is given $TC^{CNS} = E\mathrm{IC}_1^{AR} + E\mathrm{IC}_2^{AR,NS}$.

It is clear that the retailer will have a lower expected inventory cost when using optimal forecasting as opposed to exponential smoothing. It can be seen that this is due to the retailer having a larger MSFE in Equation (14) than in Equation (10) since the equations are identical otherwise.

It is not clear whether or not the supplier would benefit from the retailer using optimal forecasting. Furthermore, if the supplier shares its forecasting expertise with the retailer, the supplier may not be able to recover the retailer's demand shocks (see Remark 2 in Section 4). Finally, even if the supplier does not directly benefit from the retailer's use of optimal forecasting, the chain as a whole may be better off (the retailer and supplier can split the reduction in cost).

To better understand how the forecasting approach used by the retailer can affect the cost structure of the supply chain, we consider a supply chain where the retailer and supplier both have unit holding and shortage costs of 1 and 9 respectively. Below, we provide graphs that help demonstrate the benefits and/or consequences of the supplier sharing its forecasting expertise so the retailer can use optimal forecasting.

In Figure 2, we graph the ratio of the retailer's expected cost under exponential smoothing to optimal forecasting where $\alpha = .45$ and $\ell_1 = 6$. This graph demonstrates that the retailer's expected cost can be more than 10 times higher when it uses exponential smoothing as opposed to optimal forecasting. In Figure 3, we show the regions where the supplier's expected cost is higher if the retailer uses exponential smoothing as opposed to optimal forecasting. It can be seen that whenever $\phi < \alpha$, the supplier is benefited by the retailer using optimal forecasting. For larger values of ℓ_2, the region where the supplier has a higher expected cost under the retailer using optimal forecasting (as opposed to exponential smoothing) becomes smaller.

In Figure 4, we show the regions where the chain as a whole is better off under the retailer using optimal forecasting as opposed to exponential smoothing. In Figure 4, we assume that the retailer has shared its demand shocks when beneficial to the supplier. It can be seen that whenever $\phi < \alpha$, the chain is benefited by the retailer using optimal forecasting. For larger values of ℓ_2, the region when the chain has a higher expected cost under the retailer using optimal forecasting (as opposed to exponential smoothing) becomes smaller. When ϕ is close to 1 and α is close to 0, the chain is also better off when the retailer uses optimal forecasting.

In Figure 5, the green shaded region is where the supplier is better off under the retailer using optimal forecasting and hence the chain as a whole is better off. The red shaded region is where the chain as a whole is better off under the retailer using optimal forecasting although the supplier is better off when the retailer uses exponential smoothing. The blue shaded region is where the chain is worse off when the retailer uses optimal forecasting as opposed to exponential smoothing. In Figure 5, we assume that the retailer has shared its demand shocks where beneficial to the supplier.

In Figure 6, we graph the ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and shares its demand shocks when $\alpha = .3$ and $\ell_2 = 0$. This demonstrates that when the retailer uses exponential smoothing as opposed to the retailer using optimal forecasting and sharing its demand shocks with the supplier, the supplier's expected cost can be more than 100 times higher. Figure 7 includes a similar graph compared to Figure 6 except that in Figure 7, the retailer does not share its demand shocks. When $\phi \leq -.5$, the supplier is unable to recover the retailer's demand shocks (see Proposition 11) and hence, there is less benefit to the supplier providing forecasting expertise to the retailer when the retailer will not share its demand shocks.

In Figure 8, we provide a surface plot of the ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and shares its demand shocks (when beneficial to the supplier) versus α and ϕ where the leadtime is 0. Figure 9 includes a similar surface plot, except here, the leadtime is 6. These plots demonstrate that the supplier is most benefited by sharing its forecasting expertise with the retailer when the retailer in return shares its demand shocks with the supplier. It can be seen from Figures 8 and 9, that the ratio of the supplier's cost where the retailer uses exponential smoothing to where the retailer uses optimal forecasting and shares its demand shocks (when beneficial to the supplier) is largest when the retailer's sharing of its demand shocks is indeed beneficial to the supplier. In Proposition 11, we show where the supplier is unable to recover the retailer's demand shocks (i.e., the retailer's order to the supplier is not invertible with respect to the retailer's demand shocks).

4 Forecasting and Propagation Results

In this section, we provide theoretical results on the retailer's order process based on its optimal and suboptimal forecast of demand. We then derive the MSFE of the leadtime demand for both the retailer and the supplier under the retailer's optimal and suboptimal forecast. We also provide sufficient conditions under which the supplier is able to infer the retailer's demand shocks when the retailer adopts the optimal forecasting method.

4.1 The Retailer

If the retailer uses exponential smoothing, its forecast of its AR(1) demand at time $t+1$ based on information available at t denoted by $\hat{D}^{ES}_{1,t+1}$ is

$$\hat{D}^{ES}_{1,t+1} = \alpha \sum_{k=1}^{\infty} (1-\alpha)^{k-1} D_{1,t+1-k} \tag{16}$$

where $0 < \alpha < 1$ is the smoothing parameter. The exponential smoothing method uses the present and all past observations and assigns a weight for each observation, where the current observation receives the highest weight. If the retailer's true demand process is $ARIMA(0,1,1)$, the exponential smoothing forecast (with $\alpha = 1-\theta$, where θ is the moving average coefficient in the ARIMA(0,1,1) model) is its best liner forecast at time $t+1$. Since the retailer's true demand process is $AR(1)$, the forecast generated by the exponential smoothing is not the best linear forecast.

Proposition 5 *The retailer's forecast over its leadtime based on the exponential smoothing method* *is given by*

$$m_{1,t}^{ES} = \sum_{i=1}^{\ell_1+1} \hat{D}_{1,t+i}^{ES} = (\ell_1 + 1)\left[\frac{\alpha}{1-(1-\alpha)B}\right] D_{1,t}. \tag{17}$$

Proofs of Proposition 5 as well as subsequent propositions are provided in the Appendix.

Below we show that the retailer's order process is an ARMA(2,1) process.

Proposition 6 *The retailer's order based on the exponential smoothing method is ARMA(2,1) and is given by*

$$[1-(1-\alpha)B]\,(1-\phi B) D_{2,t}^{ES} = \alpha d + \epsilon_{1,t}^{ES} - \left(\frac{1+\alpha\ell_1}{1+\alpha+\alpha\ell_1}\right)\epsilon_{1,t-1}^{ES} \tag{18}$$

where $\epsilon_{1,t}^{ES} = (1+\alpha+\alpha\ell_1)\epsilon_{1,t}^{true}$.

Remark 1 *Since* $0 < \alpha < 1$*, the root of the polynomial* $\theta(z) = 1 - \left(\frac{1+\alpha\ell_1}{1+\alpha+\alpha\ell_1}\right) z$ *is outside the unit circle. Therefore, the retailer's order to the supplier* $\left\{D_{2,t}^{ES}\right\}$ *is invertible with respect to the retailer's true demand shocks* $\{\epsilon_{1,t}^{true}\}$*. In other words, the retailer's current demand shock can be obtained as a linear combination of the retailer's present and past order observations. Therefore, the supplier can recover the retailer's true demand shocks and utilize them to forecast its leadtime demand. Note that the roots of the polynomial* $\phi(z) = [1-(1-\alpha)z]\,(1-\phi z)$ *are both outside the unit circle. Hence the retailer's order process is causal.*

Proposition 7 *The retailer's MSFE based on the exponential smoothing method is equal to*

$$\nu_1^{ES} = E\left(\sum_{i=1}^{l_1+1} D_{1,t+i} - m_{1,t}^{ES}\right)^2 = \sum_{k=0}^{\infty} \lambda_{1,k}^2 \sigma_{\epsilon_1}^2 = \left(\sum_{k=0}^{\ell_1} \lambda_{1,k}^2 + \sum_{k=\ell_1+1}^{\infty} \lambda_{1,k}^2\right)\sigma_{\epsilon_1}^2 \tag{19}$$

where the $\lambda_{1,k}$ *are defined in Equations (39), (40), and (41) in the Appendix, and* $\sum_{k=\ell_1+1}^{\infty} \lambda_{1,k}^2 \sigma_{\epsilon_1}^2$ *has an analytical expression.*

On the other hand, if the supplier provides the retailer with forecasting expertise, the retailer can forecast its AR(1) demand optimally. Under such circumstances, the retailer uses its best linear forecast of demand over the leadtime. We summarize the retailer's forecast, order process, and MSFE in the following propositions. These results follow from GHS (2014), and are included to show the difference between optimal and suboptimal forecasting performance.

Proposition 8 *The retailer's best linear forecast of its demand over the leadtime is*

$$m_{1,t}^{AR} = d\sum_{k=0}^{\ell_1}\sum_{j=0}^{k}\phi^j + \sum_{k=1}^{\ell_1+1}\phi^k D_{1,t}.$$

The retailer's order process based upon the best linear forecast $\{D_{2,t}^{AR}\}$ *is an* $ARMA(1,1)$ *process*

$$(1-\phi B)D_{2,t}^{AR} = d + \epsilon_{1,t}^{AR,S} - \left[\frac{\phi(1-\phi^{\ell_1+1})}{1-\phi^{\ell_1+2}}\right]\epsilon_{1,t-1}^{AR,S} \quad (20)$$

where $\epsilon_{1,t}^{AR,S} = \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)\epsilon_{1,t}^{true}$.

Remark 2 *The retailer's order process* $\{D_{2,t}^{AR}\}$ *may not be invertible with respect to the retailer's true shocks, since the root of the polynomial* $\tilde{\theta}(z) = 1 - \left[\frac{\phi(1-\phi^{\ell_1+1})}{1-\phi^{\ell_1+2}}\right]z$ *may be inside or outside the unit circle, depending on the sign of the* $AR(1)$ *coefficient* ϕ *and whether* ℓ_1 *is odd or even (see Proposition 11).*

Proposition 9 *The MSFE based upon the retailer's best linear forecast is given by*

$$\nu_1^{AR} = E\left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{AR}\right)^2 = \sum_{k=0}^{\ell_1}\omega_{1,k}^2\sigma_{\epsilon_1}^2, \quad (21)$$

where the $\omega_{1,k}$ *are defined in Equation (50) in the Appendix.*

Remark 3 *We note that the retailer's MSFE under the correct AR(1) model only includes the variance of demand shocks occurring during the forecasting period, while the MSFE under the exponential smoothing forecast includes the variance of shocks occurring during the leadtime as well as the infinite past.*

4.2 The Supplier

Unlike the retailer, the supplier always uses the best linear forecast of its leadtime demand. However, as shown above, the retailer's order and hence the demand that the supplier must serve depends upon the forecasting methodology used by the retailer.

If the retailer uses the exponential smoothing forecasting methodology, then the supplier faces ARMA(2,1) demand, $\{D_{2,t}^{ES}\}$. Next we provide a proposition describing the supplier's best linear forecast and its associated MSFE.

Proposition 10 *If the retailer uses the exponential smoothing forecasting methodology, the supplier's best linear forecast of demand over the leadtime is*

$$m_{2,t}^{ES} = \sum_{k=\ell_2+1}^{\infty} \xi_k^{ES} \epsilon_{1,t+\ell_2+1-k}^{ES}$$

and its associated MSFE is

$$\nu_2^{ES} = E\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{ES} - m_{2,t}^{ES}\right)^2 = [1+\alpha(1+\ell_1)]^2 \sum_{k=0}^{\ell_2} (\xi_k^{ES})^2 \sigma_{\epsilon_1}^2,$$

where the ξ_k^{ES} are defined in Equation (60), (52), (53), and (54) in the Appendix.

If the retailer uses optimal forecasting, then if the root of the polynomial $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell 1+2}}\right) z$ is outside the unit circle, i.e. $\left|\frac{1-\phi^{\ell 1+2}}{\phi-\phi^{\ell_1+2}}\right| > 1$, the supplier's demand is invertible with respect to the retailer's demand shocks $\{\epsilon_{1,t}^{true}\}$. In other words, in such a case, the supplier can recover the retailer's demand shocks from the supplier's own demand. On the other hand, if the root of the polynomial $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell 1+2}}\right) z$ is inside the unit circle, i.e. $\left|\frac{1-\phi^{\ell 1+2}}{\phi-\phi^{\ell_1+2}}\right| < 1$, the supplier's demand is not invertible with respect to the retailer's demand shocks $\{\epsilon_{1,t}^{true}\}$. In such a case, the supplier cannot recover the retailer's demand shocks from the supplier's own demand. However, the retailer may or may not share its demand shocks with the supplier in such an instance. The propositions below summarize the sufficient conditions under which the supplier's demand is invertible with respect to the retailer's demand shocks, and the supplier's best linear forecast and its associated MSFE under the retailer's sharing and non-sharing of its demand shocks.

Proposition 11 *Suppose the retailer's demand is an AR(1) process with AR coefficient* $\phi \in (-1,1)$*. If* $\phi \in (0,1)$*, the supplier's demand is always invertible with respect to the retailer's demand shocks* $\{\epsilon_{1,t}^{true}\}$*. If* $\phi \in (-1,0)$*, then*

i) if ℓ_1 *is odd, the supplier's demand is invertible with respect to the retailer's demand shocks* $\{\epsilon_{1,t}^{true}\}$*.*

ii) if ℓ_1 *is even, then there exists a constant* $\kappa(\ell_1) \in (-1,0)$ *such that if* $\phi \in (-1, \kappa(\ell_1))$*, the supplier's demand is not invertible with respect to the retailer's demand shocks* $\{\epsilon_{1,t}^{true}\}$*. If* $\phi \in [\kappa(\ell_1), 0)$*, the supplier's demand is invertible with respect to the retailer's demand shocks* $\{\epsilon_{1,t}^{true}\}$*.*

Proposition 12 *If the retailer uses optimal forecasting and the root of the polynomial* $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell 1+2}}\right) z$ *is outside the unit circle, or the root of the polynomial* $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell 1+2}}\right) z$ *is inside the unit circle but the retailer shares its demand shocks with the supplier, the supplier's best linear forecast of its demand over the leadtime is given by*

$$m_{2,t}^{AR,S} = (1+\ell_2)\mu_d + \sum_{k=\ell_2+1}^{\infty} \xi_k^{AR} \epsilon_{1,t+\ell_2+1-k}^{AR,S}$$

and its associated MSFE is

$$\nu_2^{AR,S} = E\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} - m_{2,t}^{AR,S}\right)^2 = \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)^2 \sum_{k=0}^{\ell_2} (\xi_k^{AR})^2 \sigma_{\epsilon_1}^2$$

where the ξ_k^{AR} *are defined in Equations (63) and (64) in the Appendix.*

Proposition 13 *If the retailer uses optimal forecasting and the root of the polynomial* $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell 1+2}}\right) z$ *is inside the unit circle and the retailer does not share its demand shocks with the supplier, the supplier's best linear forecast of its demand over the leadtime is given by*

$$m_{2,t}^{AR,NS} = (1+\ell_2)\mu_d + \sum_{k=\ell_2+1}^{\infty} \xi_{2,k} \tilde{\epsilon}_{2,t+\ell_2+1-k}$$

where

$$\tilde{\epsilon}_{2,t} = \left[\frac{1}{1 - \left(\frac{1-\phi^{\ell_1+2}}{\phi-\phi^{\ell_1+2}}\right) B} \right] \left[(1-\phi B) D_{2,t}^{AR} - d \right]$$

and the MSFE of the supplier's best linear forecast is equal to

$$\nu_2^{AR,NS} = E\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} - m_{2,t}^{AR,NS} \right)^2 = \left(\frac{\phi - \phi^{\ell_1+2}}{1-\phi} \right)^2 \sum_{k=0}^{\ell_2} (\xi_{2,k})^2 \sigma_{\epsilon_1}^2$$

where the $\xi_{2,k}$ *are defined in Equations (69) and (70) in the Appendix.*

5 Conclusion

Although it is common for firms to forecast stationary demand using simple exponential smoothing due to the ease of computation and understanding of the methodology, we have shown that the costs of doing so can be significant. Indeed, we have shown that a retailer using exponential smoothing may have expected inventory related costs more than ten times higher than when compared to using the optimal forecast. We demonstrated that when $\phi < \alpha$, the suppliers expected inventory related cost is less when the retailer uses optimal forecasting as opposed to exponential smoothing. We have also shown that there exists an additional set of cases where the sum of the expected inventory related costs of the retailer and the supplier is less when the retailer uses optimal forecasting as opposed to exponential smoothing even though the supplier's expected costs are higher.

This research dovetails nicely with other information sharing papers as well. We demonstrated that the supplier has most to gain by sharing its forecasting expertise with the retailer when the retailer's order to the supplier is not invertible with respect to the retailer's demand shocks. Indeed, we have shown that the the supplier's expected cost where the retailer uses exponential smoothing can be more than 100 times the supplier's expected cost where the retailer uses optimal forecasting and shares its demand shocks.

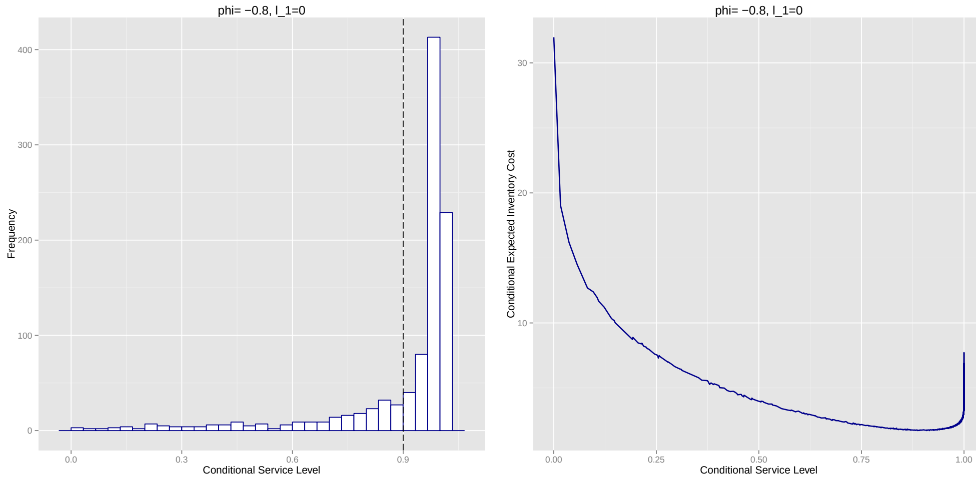

Figure 1: Top: histogram of the retailer's conditional service level $\Phi(r)$. Bottom: retailer's conditional expected inventory cost vs. conditional service level. The optimal service level, $\frac{s}{s+h} = .9$ as $s = 9$ and $h = 1$.

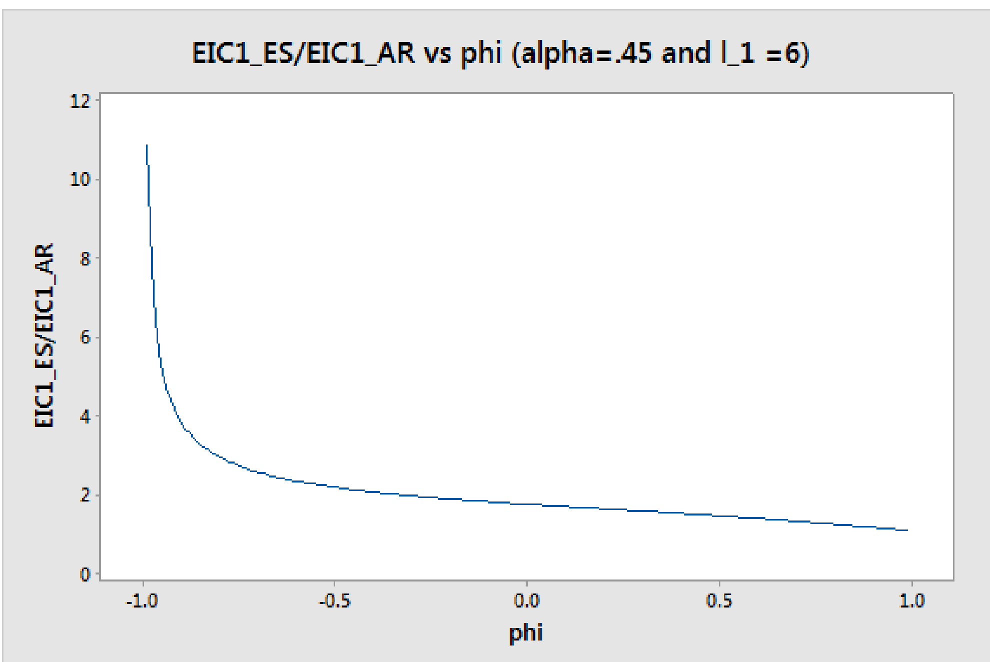

Figure 2: The ratio of the retailer's expected cost when it uses exponential smoothing as opposed to optimal forecasting versus ϕ. Here $\alpha = .45$ and $\ell_1 = 6$.

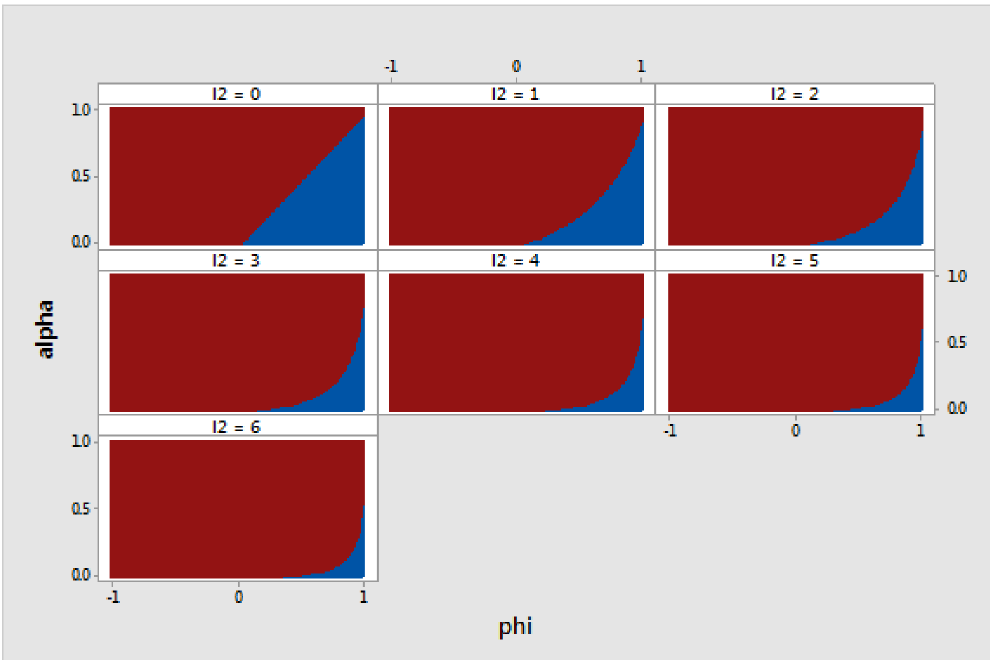

Figure 3: Graph of α versus ϕ where ℓ_2 varies from 0 to 6. The red shaded region is where $EIC_2^{ES} \geq EIC_2^{AR}$. The blue shaded region is where $EIC_2^{ES} < EIC_2^{AR}$. It is clear that for most values of α and ϕ, the supplier has a lower expected cost when the retailer uses optimal forecasting.

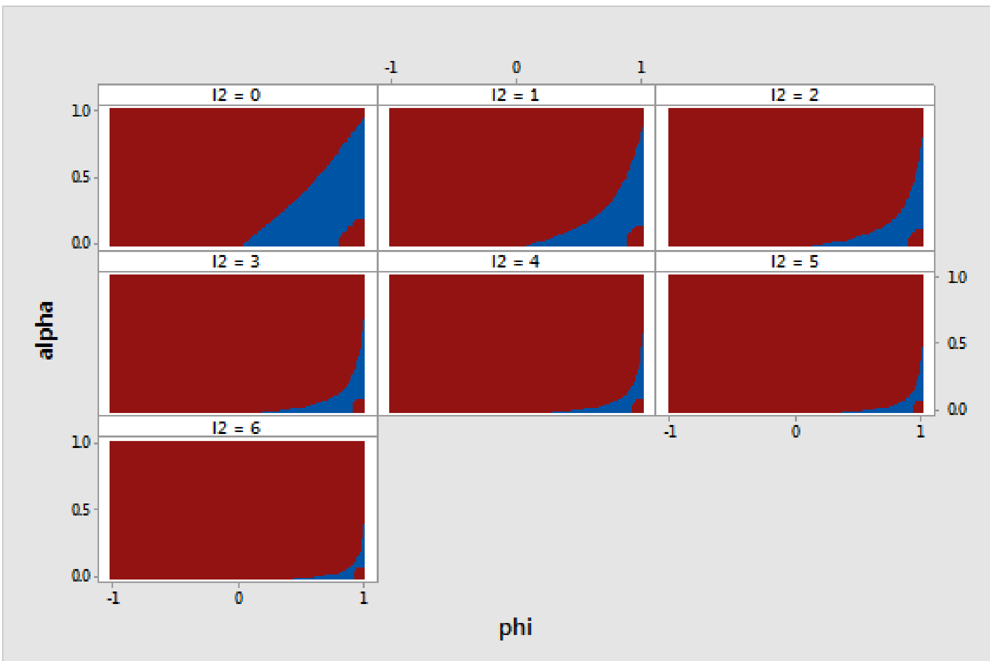

Figure 4: Graph of α versus ϕ where ℓ_2 varies from 0 to 6. The red shaded region is where $TC^{ES} \geq TC^{AR}$. The blue shaded region is where $TC^{ES} < TC^{AR}$. It is clear that for most values of α and ϕ, the chain has a lower expected cost when the retailer uses optimal forecasting.

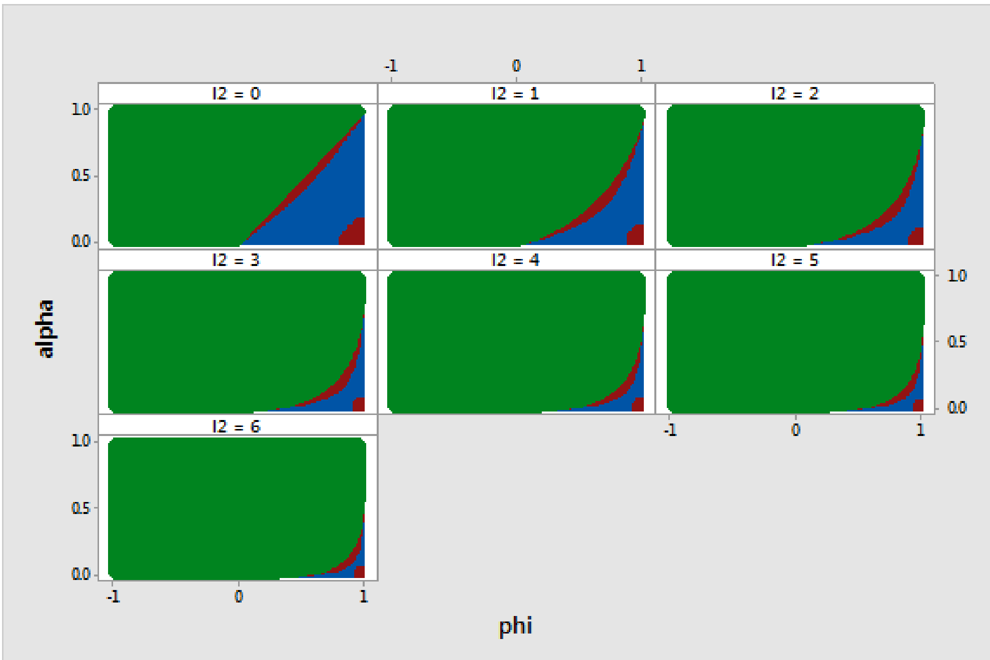

Figure 5: The green shaded region is where the supplier is better off under the retailer using optimal forecasting and hence the chain as a whole is better off. The red shaded region is where the chain as a whole is better off under the retailer using optimal forecasting although the supplier is better off when the retailer uses exponential smoothing. The blue shaded region is where the chain is worse off when the retailer uses optimal forecasting as opposed to exponential smoothing.

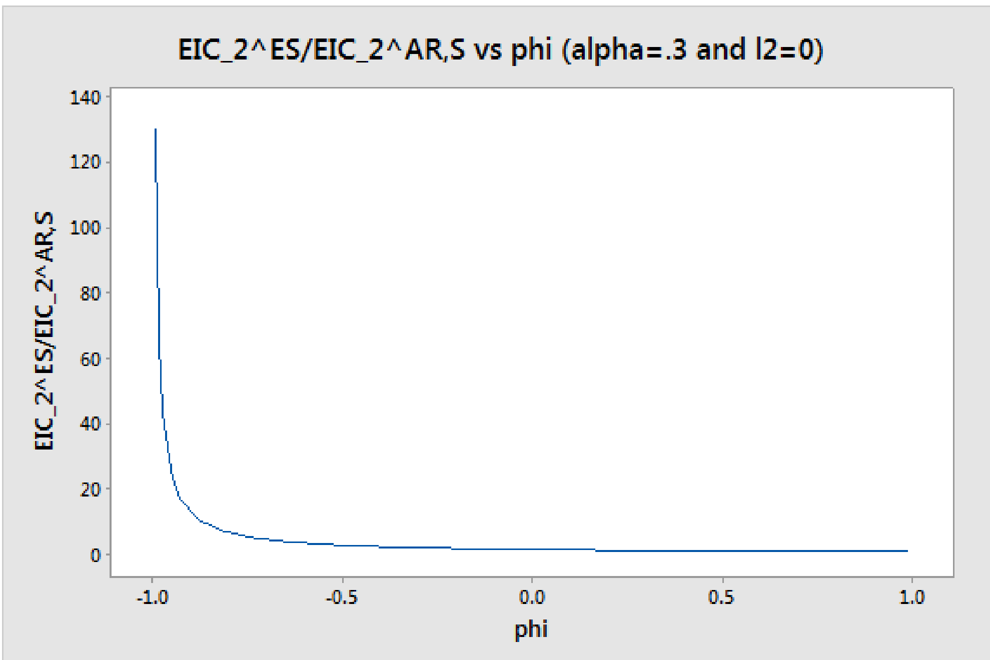

Figure 6: The ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and shares its demand shocks when $\alpha = .3$ and $\ell_2 = 0$.

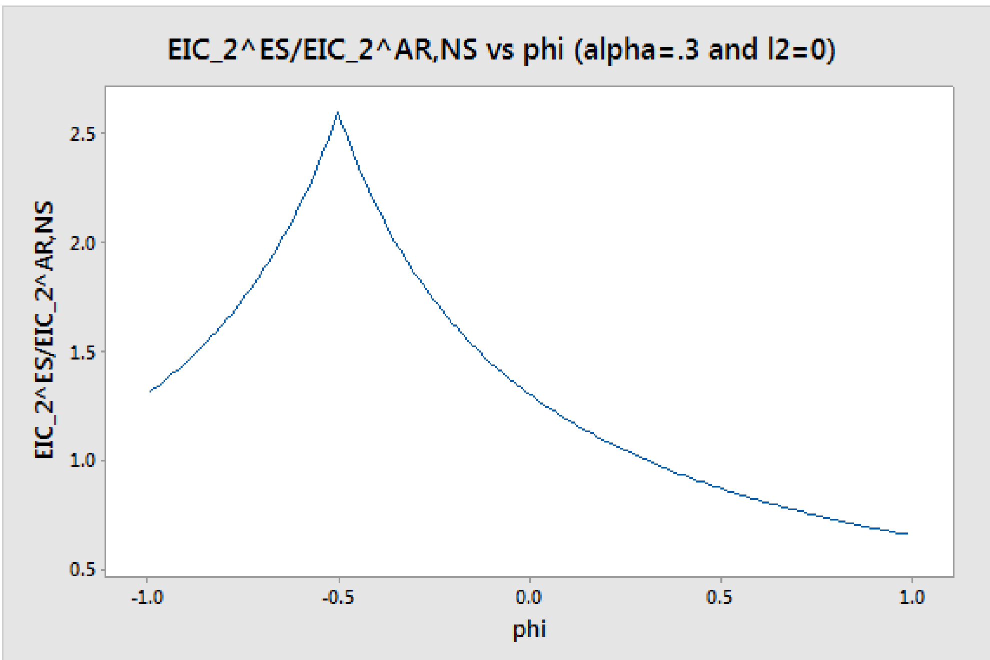

Figure 7: The ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and does not share its demand shocks when $\alpha = .3$ and $\ell_2 = 0$.

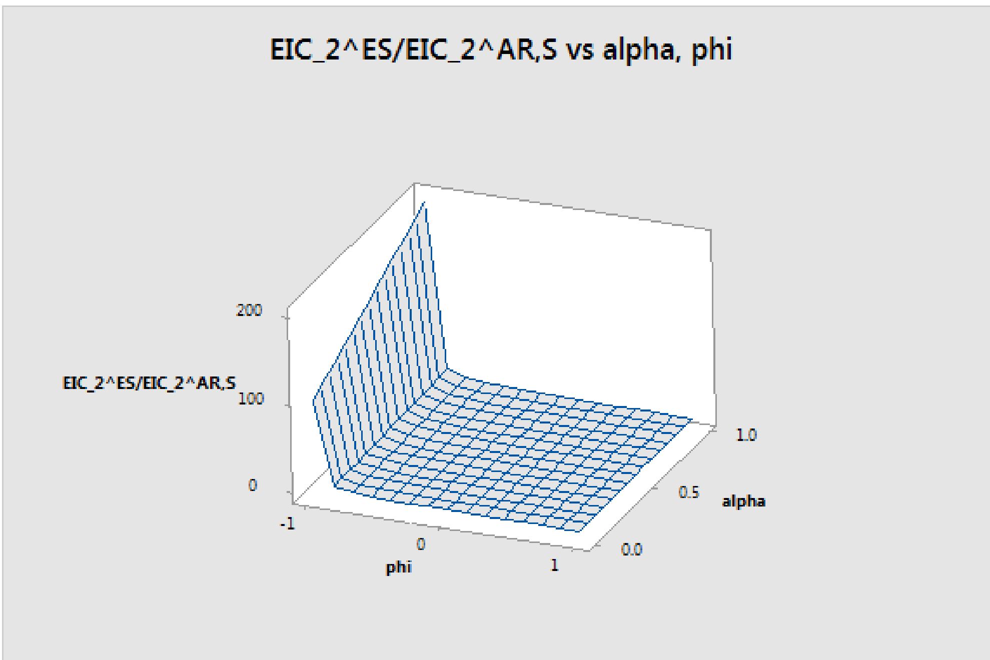

Figure 8: The ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and shares its demand shocks where $\ell_1 = \ell_2 = 0$.

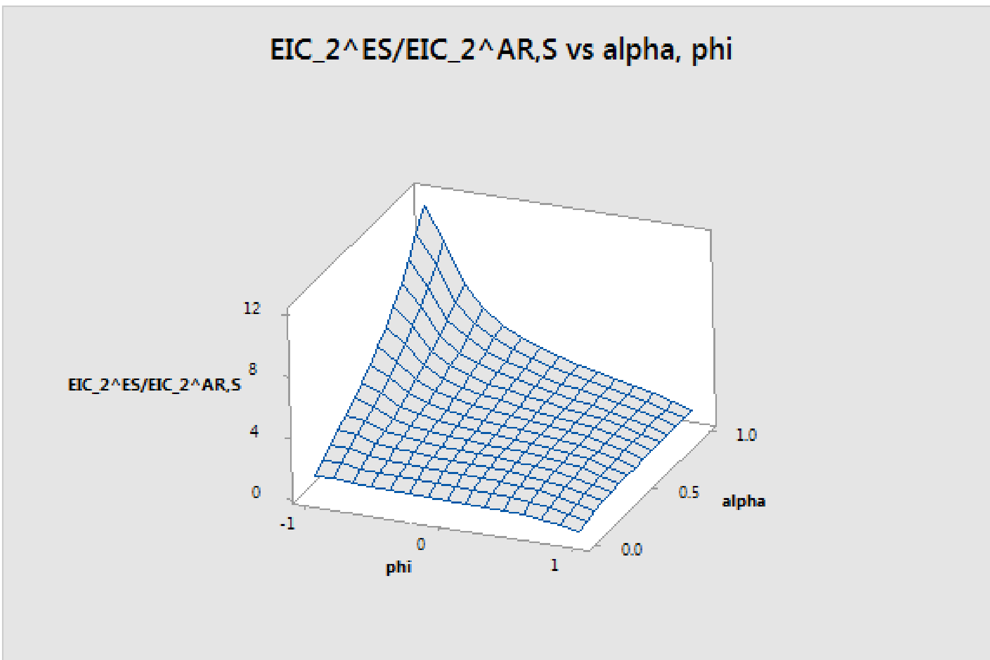

Figure 9: The ratio of the supplier's cost when the retailer uses exponential smoothing to when the retailer uses optimal forecasting and shares its demand shocks where $\ell_1 = \ell_2 = 6$.

Appendix

Proof of Proposition 1:

Let $X = \sum_{i=1}^{l_1+1} D_{1,t+i}$. Then the conditional distribution of X given $\mathcal{M}_t^1$ is normal with mean $m_{1,t}^*$ and variance ν_1^*. Let F be the CDF of this conditional distribution. The retailer's conditional expected cost under the optimal policy can be expressed as

$$s\int_{S_{1,t}^*}^{\infty}(x - S_{1,t}^*)dF(x) - h\int_{-\infty}^{S_{1,t}^*}(x - S_{1,t}^*)dF(x). \tag{22}$$

Let $Z = \frac{X - m_{1,t}^*}{\sqrt{\nu_1^*}}$ so that $X - S_{1,t}^* = X - m_{1,t}^* - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^*} = \sqrt{\nu_1^*}\left(Z - \Phi^{-1}\left(\frac{s}{s+h}\right)\right)$ and $dF(X) = d\Phi(Z)$. Thus, we can express (22) as

$$\sqrt{\nu_1^*}[s\int_{\Phi^{-1}\left(\frac{s}{s+h}\right)}^{\infty}\left(z - \Phi^{-1}\left(\frac{s}{s+h}\right)\right)d\Phi(z) - h\int_{-\infty}^{\Phi^{-1}\left(\frac{s}{s+h}\right)}\left(z - \Phi^{-1}\left(\frac{s}{s+h}\right)\right)d\Phi(z)]. \tag{23}$$

From Equation (9), note that $\int_{-\infty}^{x}(z-x)d\Phi(z) = \int_{-\infty}^{\infty}(z-x)d\Phi(z) - L(x) = -x - L(x)$. Therefore, Equation (23) becomes

$$\sqrt{\nu_1^*}\left[(s+h)L\left(\Phi^{-1}\left(\frac{s}{s+h}\right)\right) + h\Phi^{-1}\left(\frac{s}{s+h}\right)\right] \tag{24}$$

which is the conditional expected cost of the retailer's cost under the optimal policy (see Equation 4.7 in LST). □

Proof of Proposition 3:

The retailer's conditional expected cost under exponential smoothing can be expressed as

$$s\int_{S_{1,t}^{ES}}^{\infty}(x - S_{1,t}^{ES})dF(x)|\mathcal{M}_t^1) - h\int_{-\infty}^{S_{1,t}^{ES}}(x - S_{1,t}^{ES})dF(x)|\mathcal{M}_t^1). \tag{25}$$

Since $Z = \frac{X - m_{1,t}^*}{\sqrt{\nu_1^*}}$, it follows that

$$X - S_{1,t}^{ES} = X - m_{1,t}^{ES} - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}} = \sqrt{\nu_1^*}Z + m_{1,t}^* - m_{1,t}^{ES} - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}. \tag{26}$$

Therefore the retailer's conditional expected cost is

$$\sqrt{\nu_1^*}\left[s\int_{r}^{\infty}(z-r)d\Phi(z) - h\int_{-\infty}^{r}(z-r)d\Phi(z)\right]. \tag{27}$$

It follows that Equation (27) representing the retailer's conditional expectation of cost under exponential smoothing is random and a function of r, where $\Phi(r)$ is the conditional expected service level, which is also a random variable. $\square$

Proof of Proposition 4:

The retailer's forecast error under the exponential smoothing is

$$\sum_{i=1}^{l_1+1} D_{1,t+i} - m_{1,t}^{ES} = \sum_{k=0}^{\infty} \lambda_{1,k}\epsilon_{1,t+\ell_1+1-k}^{true} \tag{28}$$

Let $Y = \sum_{i=1}^{\ell_1+1} D_{t+i} - m_{1,t}^{ES}$. Then the unconditional distribution of Y is normal with mean zero and variance ν_1^{ES}. The retailer's unconditional expected cost under exponential smoothing can be written as

$$E\left\{\left[\left(\sum_{i=1}^{\ell_1+1} D_{t+i} - S_{1,t}^{ES}\right)^+ s + \left(\sum_{i=1}^{\ell_1+1} D_{t+i} - S_{1,t}^{ES}\right)^- h\right]\right\} \tag{29}$$

$$= E\left\{\left[\left(\sum_{i=1}^{\ell_1+1} D_{t+i} - m_{1,t}^{ES} - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right)^+ s\right.\right.$$

$$+ \left.\left.\left(\sum_{i=1}^{\ell_1+1} D_{t+i} - m_{1,t}^{ES} - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right)^- h\right]\right\}$$

$$= E\left[\left(Y - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right)^+ s + \left(Y - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right)^- h\right] \tag{30}$$

which can also be expressed as

$$s\int_{\Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}}^{\infty} \left(y - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right) dG(y)$$

$$- \; h\int_{-\infty}^{\Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}} \left(y - \Phi^{-1}\left(\frac{s}{s+h}\right)\sqrt{\nu_1^{ES}}\right) dG(y) \tag{31}$$

where $G(\cdot)$ is the CDF of Y.

Let $Z = \frac{Y}{\sqrt{\nu_1^{ES}}} = \frac{\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{ES}}{\sqrt{\nu_1^{ES}}}$ so that $Y = Z\sqrt{\nu_1^{ES}}$ and $dG(Y) = d\Phi(Z)$. After the transformation of Y, Equation (31) becomes

$$\sqrt{\nu_1^{ES}}\left[s\int_{\Phi^{-1}\left(\frac{s}{s+h}\right)}^{\infty}\left(z - \Phi^{-1}\left(\frac{s}{s+h}\right)\right)d\Phi(z) - h\int_{-\infty}^{\Phi^{-1}\left(\frac{s}{s+h}\right)}\left(z - \Phi^{-1}\left(\frac{s}{s+h}\right)\right)d\Phi(z)\right] \tag{32}$$

note that $\int_{-\infty}^{x}(z-x)d\phi(z) = \int_{-\infty}^{\infty}(z-x)d\phi(z) - L(x) = -x - L(x)$. Therefore, Equation (32) is equal to

$$\sqrt{\nu_1^{ES}}\left[(s+h)L\left(\Phi^{-1}\left(\frac{s}{s+h}\right)\right) + h\Phi^{-1}(\frac{s}{s+h})\right]. \tag{33}$$

□

Proof of Proposition 5.

Define $\Lambda(B) = 1 + (1-\alpha)B + (1-\alpha)^2B^2 + (1-\alpha)^3B^3 + \ldots = \frac{\alpha}{1-(1-\alpha)B}$, where B is the backshift operator such that $BD_{1,t} = D_{1,t-1}$. Then equation (16) can be represented as

$$\hat{D}_{1,t+1}^{ES} = \alpha\Lambda(B)D_{1,t} = \left[\frac{\alpha}{1-(1-\alpha)B}\right]D_{1,t} \tag{34}$$

The h-steps ahead forecast of the retailer's demand based on retailer's available information set at time t $\mathcal{M}_t^1$, $\hat{D}_{t+h}^{ES}$, is the same as $\hat{D}_{t+1}^{ES}$. Therefore, the linear forecast of the future demand during the replenishment period ℓ_1 can be expressed as

$$m_{1,t}^{ES} = \sum_{i=1}^{\ell_1+1}\hat{D}_{1,t+i}^{ES} = (\ell_1+1)\hat{D}_{1,t+1}^{ES} = (\ell_1+1)\left[\frac{\alpha}{1-(1-\alpha)B}\right]D_{1,t}. \tag{35}$$

□

Proof of Proposition 6:

Let $\tilde{D}_{1,t} = D_{1,t} - \mu_d$ be the demeaned demand process, where $\mu_d = E\left[D_{1,t}\right] = \frac{d}{1-\phi}$. Then

$$\tilde{D}_{1,t} = (1-\phi B)^{-1}\epsilon_{1,t}^{true} \tag{36}$$

and

$$m_{1,t}^{ES} = (\ell_1+1)\alpha\Lambda(B)\tilde{D}_{1,t} + (1+\ell_1)\mu_d. \tag{37}$$

Using the results from Equations (36) and (37), one can represent the retailer's order with respect to the retailer's true demand shocks as

$$
\begin{aligned}
D_{2,t}^{ES} &= D_{1,t} + (1-B)m_{1,t}^{ES} \\
&= \mu_d + \left[1 + (1-B)\alpha\Lambda(B)(1+\ell_1)\right]\tilde{D}_{1,t} \\
&= \mu_d + \frac{(1-B)(1+\alpha\ell_1)+\alpha}{\left[1-(1-\alpha)B\right](1-\phi B)}\epsilon_{1,t}^{true}. \qquad (38)
\end{aligned}
$$

Applying the operator $\left[1-(1-\alpha)B\right](1-\phi B)$ to both sides of (38), it follows that

$$
\left[1-(1-\alpha)B\right](1-\phi B)D_{2,t}^{ES} = \alpha d + \left[\alpha + (1-B)(1+\alpha\ell_1)\right]\epsilon_{1,t}^{true}.
$$

Let $\epsilon_{1,t}^{ES} = (1+\alpha+\alpha\ell_1)\epsilon_{1,t}^{true}$. Then Equation (38) can be expressed as

$$
\left[1-(1-\alpha)B\right](1-\phi B)D_{2,t}^{ES} = \alpha d + \epsilon_{1,t}^{ES} - \left(\frac{1+\alpha\ell_1}{1+\alpha+\alpha\ell_1}\right)\epsilon_{1,t-1}^{ES}.
$$

$\square$

Lemma 1 *The retailer's forecast errors over the leadtime under the suboptimal forecast are equal to*

$$
\sum_{i=1}^{l_1+1} D_{1,t+i} - m_{1,t}^{ES} = \sum_{k=0}^{\infty} \lambda_{1,k}\epsilon_{1,t+\ell_1+1-k}^{true}
$$

where

$$
\lambda_{1,k} = \begin{cases} 1 & : \quad k = 0 \\ \lambda_{1,k-1} + \phi^k & : \quad 0 < k \le \ell_1 \\ \sum_{s=1}^{\ell_1+1} \psi_{s,s+k-\ell_1-1} & : \quad k > \ell_1 \end{cases} \qquad (39)
$$

with

$$
\psi_{i,j} = \begin{cases} \phi^j & : \quad 0 \le j \le i-1 \\ \left(\phi^i - \frac{\alpha\phi}{\alpha+\phi-1}\right)\phi^{j-i} - \frac{\alpha(\alpha-1)}{\alpha+\phi-1}(1-\alpha)^{j-i} & : \quad i \le j \end{cases} \qquad (40)
$$

if $\alpha + \phi \neq 1$

and

$$
\psi_{i,j} = \begin{cases} \phi^j & : \quad 0 \le j \le i-1 \\ \left[\phi^i - \alpha(j-i+1)\right]\phi^{j-i} & : \quad i \le j \end{cases} \qquad (41)
$$

if $\alpha + \phi = 1$.

Proof of Lemma 1:

The retailer's forecast errors over the leadtime are

$$\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{ES} = \sum_{i=1}^{\ell_1+1} \left[D_{1,t+i} - \frac{m_{1,t}^{ES}}{1+\ell_1} \right]$$
$$= \sum_{i=1}^{\ell_1+1} \left[\tilde{D}_{1,t+i} + \mu_d - \frac{m_{1,t}^{ES}}{1+\ell_1} \right] \quad (42)$$

Using the expression for $m_{1,t}^{ES}$ from Equation (37), the retailer's forecast error at $t+i$ can be expressed as

$$D_{1,t+i} - \frac{m_{1,t}^{ES}}{1+\ell_1} = \tilde{D}_{1,t+i} + \mu_d - \frac{m_{1,t}^{ES}}{1+\ell_1} = \left[\tilde{D}_{1,t+i} - \alpha\Lambda(B)\tilde{D}_{1,t} \right]$$
$$= \left(\phi^i \tilde{D}_{1,t} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true} \right) - \alpha\Lambda(B)\tilde{D}_{1,t}$$
$$= \left[\phi^i - \alpha\Lambda(B) \right] \tilde{D}_{1,t} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true}$$
$$= \left[\phi^i - \frac{\alpha}{1-(1-\alpha)B} \right] (1-\phi B)^{-1} \epsilon_{1,t}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true}.$$

Case I: $\alpha + \phi \neq 1$**.**

If $\alpha + \phi \neq 1$, one can apply partial fractions to represent

$$\left[\phi^i - \frac{\alpha}{1-(1-\alpha)B} \right] (1-\phi B)^{-1} \epsilon_{1,t}^{true} = \left\{ \frac{\phi^i}{1-\phi B} - \alpha \left[\frac{\Gamma_1}{1-(1-\alpha)B} + \frac{\Gamma_2}{1-\phi B} \right] \right\} \epsilon_{1,t}^{true}$$

where $\Gamma_1 = \frac{\alpha-1}{\alpha+\phi-1}$ and $\Gamma_2 = \frac{\phi}{\alpha+\phi-1}$. Hence

$$\begin{aligned}
D_{1,t+i} - \frac{m_{1,t}^{ES}}{1+\ell_1} &= \left[\phi^i \sum_{j=0}^{\infty} \phi^j B^j - \alpha\Gamma_1 \sum_{j=0}^{\infty} (1-\alpha)^j B^j - \alpha\Gamma_2 \sum_{j=0}^{\infty} \phi^j B^j\right] \epsilon_{1,t}^{true} \\
&+ \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true} \\
&= \sum_{j=0}^{\infty} \left[\left(\phi^i - \alpha\Gamma_2\right) \phi^j - \alpha\Gamma_1 (1-\alpha)^j\right] \epsilon_{1,t-j}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true} \\
&\equiv \sum_{j=0}^{\infty} \psi_{i,j} \epsilon_{1,t+i-j}^{true}
\end{aligned}$$

where

$$\psi_{i,j} = \begin{cases} \phi^j & : \quad 0 \le j \le i-1 \\ \left(\phi^i - \frac{\alpha\phi}{\alpha+\phi-1}\right) \phi^{j-i} - \frac{\alpha(1-\alpha)}{\alpha+\phi-1} (1-\alpha)^{j-i} & : \quad i \le j \end{cases}$$

Case II: $\alpha + \phi = 1$**.**

If $\alpha + \phi = 1$, then $1 - \alpha = \phi$. The retailer's forecast error at $t + i$ can be expressed as

$$\begin{aligned}
D_{1,t+i} - \frac{m_{1,t}^{ES}}{1+\ell_1} &= \left[\phi^i - \frac{\alpha}{1-\phi B}\right] (1-\phi B)^{-1} \epsilon_{1,t}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{t+i-j}^{true} \\
&= \left[\frac{\phi^i}{1-\phi B} - \frac{\alpha}{(1-\phi B)^2}\right] \epsilon_{1,t}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{t+i-j}^{true} \\
&= \left[\phi^i \sum_{j=0}^{\infty} \phi^j B^j - \alpha \sum_{j=0}^{\infty} (j+1) \phi^j B^j\right] \epsilon_{1,t}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true} \\
&= \sum_{j=0}^{\infty} \left[\phi^i - \alpha(j+1)\right] \phi^j \epsilon_{1,t-j}^{true} + \sum_{j=0}^{i-1} \phi^j \epsilon_{1,t+i-j}^{true} \\
&\equiv \sum_{j=0}^{\infty} \psi_{i,j} \epsilon_{1,t+i-j}^{true}
\end{aligned}$$

where

$$\psi_{i,j} = \begin{cases} \phi^j & : \quad 0 \le j \le i-1 \\ \left[\phi^i - \alpha(j-i+1)\right] \phi^{j-i} & : \quad i \le j \end{cases}$$

The retailer's forecast errors over the leadtime are then equal to

$$\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{ES} = \sum_{i=1}^{\ell_1+1}\sum_{j=0}^{\infty} \psi_{i,j}\epsilon_{1,t+i-j}^{true} \tag{43}$$

$$= \sum_{k=0}^{\infty} \lambda_{1,k}\epsilon_{1,t+\ell_1+1-k}^{true} \tag{44}$$

where

$$\lambda_{1,k} = \begin{cases} 1 & : \quad k=0 \\ \lambda_{1,k-1}+\phi^k & : \quad 0<k\leq \ell_1 \\ \sum_{s=1}^{\ell_1+1}\psi_{s,s+k-\ell_1-1} & : \quad k>\ell_1 \end{cases}$$

with $\psi_{s,s+k-\ell_1-1}$ defined based on the sum of α and ϕ. □

Proof of Proposition 7:

From Equation (44) in Lemma 5, it follows directly that the retailer's MSFE is equal to

$$\nu_1^{ES} = E\left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{ES}\right)^2 = \sum_{k=0}^{\infty}\lambda_{1,k}^2\sigma_{\epsilon_1}^2$$

$$= \left(\sum_{k=0}^{\ell_1}\lambda_{1,k}^2 + \sum_{k=\ell_1+1}^{\infty}\lambda_{1,k}^2\right)\sigma_{\epsilon_1}^2.$$

The analytical expression for $\sum_{k=\ell_1+1}^{\infty}\lambda_{1,k}^2$ is dependent on the sum of α and ϕ. First we consider the case when $\alpha+\phi\neq 1$. We define $\Theta(i)=\phi^i-\alpha\Gamma_2$. Together with the definition for $\psi_{i,j}$ from equation (40) when $i\leq j$, we can express

$$\lambda_{1,\ell_1+1} = \sum_{s=1}^{\ell_1+1}\psi_{s,s} = \sum_{s=1}^{\ell_1+1}\Theta(s) - \sum_{s=1}^{\ell_1+1}\alpha\Gamma_1 = \frac{\phi-\phi^{\ell_1+2}}{1-\phi} - (\ell_1+1)\alpha\Gamma_2 - (\ell_1+1)\alpha\Gamma_1.$$

For any $q\geq 1$,

$$\lambda_{1,\ell_1+q} = \sum_{s=1}^{\ell_1+1}\psi_{s,s+q-1} = \phi^{q-1}\left[\frac{\phi-\phi^{\ell_1+2}}{1-\phi} - (\ell_1+1)\alpha\Gamma_2\right] - (1-\alpha)^{q-1}\,(\ell_1+1)\alpha\Gamma_1.$$

Define $\Gamma_3 = \left[\frac{\phi-\phi^{\ell_1+2}}{1-\phi} - (\ell_1+1)\alpha\Gamma_2\right]$ and $\Gamma_4 = (\ell_1+1)\alpha\Gamma_1$. Then

$$\begin{aligned}\sum_{k=\ell_1+1}^{\infty} \lambda_{1,k}^2 &= \sum_{q=1}^{\infty} \left[\phi^{q-1}\Gamma_3 - (1-\alpha)^{q-1}\Gamma_4\right]^2 \\ &= \sum_{q=0}^{\infty} \left[\phi^{q}\Gamma_3 - (1-\alpha)^{q}\Gamma_4\right]^2 \\ &= \Gamma_3^2 \sum_{q=0}^{\infty} \phi^{2q} - 2\Gamma_3\Gamma_4 \sum_{q=0}^{\infty} [\phi(1-\alpha)]^q + \Gamma_4^2 \sum_{q=0}^{\infty}(1-\alpha)^{2q} \\ &= \frac{\Gamma_3^2}{1-\phi^2} - 2\Gamma_3\Gamma_4\left[\frac{1}{1-\phi(1-\alpha)}\right] + \frac{\Gamma_4^2}{1-(1-\alpha)^2}.\end{aligned}$$

One can apply a similar approach to find an analytical expression for $\sum_{k=\ell_1+1}^{\infty} \lambda_{1,k}^2$ when $\alpha + \phi = 1$.

For any $q \geq 1$,

$$\lambda_{1,\ell_1+q} = \sum_{s=1}^{\ell_1+q} \psi_{s,s+q-1} = \phi^{q-1}\sum_{s=1}^{\ell_1+1}(\phi^s - \alpha q) = \phi^{q-1}\left[\frac{\phi-\phi^{\ell_1+2}}{1-\phi} - \alpha(1+\ell_1)q\right].$$

Define $\Gamma_5 = \frac{\phi-\phi^{\ell_1+2}}{1-\phi}$. Then

$$\begin{aligned}\sum_{k=\ell_1+1}^{\infty} \lambda_{1,k}^2 &= \sum_{q=1}^{\infty} \phi^{2(q-1)}\left[\frac{\phi-\phi^{\ell_1+2}}{1-\phi} - \alpha(1+\ell_1)q\right]^2 \\ &= \frac{\Gamma_5^2}{1-\phi^2} - \frac{2\alpha(1+\ell_1)\Gamma_5}{(1-\phi^2)^2} + \frac{\alpha^2(\ell_1+1)^2(1+\phi^2)}{(1-\phi^2)^3},\end{aligned}$$

□

Proof of Proposition 8:

The retailer's k-step ahead forecast of demand given the available information at t is given by

$$\hat{D}_{1,t+k}^{AR} = d\sum_{j=0}^{k-1}\phi^j + \phi^k D_{1,t}. \tag{45}$$

Therefore, the retailer's best linear forecast of demand over the leadtime is

$$m_{1,t}^{AR} = \sum_{k=1}^{\ell_1+1}\hat{D}_{1,t+k}^{AR} = d\sum_{k=1}^{\ell_1+1}\sum_{j=0}^{k-1}\phi^j + \sum_{k=1}^{\ell_1+1}\phi^k D_{1,t} = d\sum_{k=0}^{\ell_1}\sum_{j=0}^{k}\phi^j + \sum_{k=1}^{\ell_1+1}\phi^k D_{1,t}.$$

One can further show that

$$m_{1,t}^{AR} - m_{1,t-1}^{AR} = \phi \left[\frac{1-\phi^{\ell_1+1}}{1-\phi} \right] (1-B) D_{1,t}$$

Hence the retailer's order process under the optimal forecast is

$$\begin{aligned} D_{2,t}^{AR} &= D_{1,t} + m_{1,t}^{AR} - m_{1,t-1}^{AR} \\ &= \mu_d + \left[\frac{1 + \phi \left(\frac{1-\phi^{\ell_1+1}}{1-\phi} \right) (1-B)}{1-\phi B} \right] \epsilon_{1,t}^{true}. \end{aligned} \tag{46}$$

Applying $1-\phi B$ to both sides of the above equation, it follows that

$$\begin{aligned} (1-\phi B) D_{2,t}^{AR} &= d + \left[1 + \phi \left(\frac{1-\phi^{\ell_1+1}}{1-\phi} \right) (1-B) \right] \epsilon_{1,t}^{true} \\ &= d + \left[1 + (1-B) \sum_{j=1}^{\ell_1+1} \phi^j \right] \epsilon_{1,t}^{true}. \end{aligned} \tag{47}$$

Rescaling the RHS of Equation (47) so that the leading MA coefficient is one, one obtains the retailer's order process

$$(1-\phi B) D_{2,t}^{AR} = d + \epsilon_{1,t}^{AR,S} - \left[\frac{\phi(1-\phi^{\ell_1+1})}{1-\phi^{\ell_1+2}} \right] \epsilon_{1,t-1}^{AR,S} \tag{48}$$

where $\epsilon_{1,t}^{AR,S} = \left(\frac{1-\phi^{\ell_1+2}}{1-\phi} \right) \epsilon_{1,t}^{true}$. □

Proof of Proposition 9:

The retailer's demand has an MA(∞) representation with respect to its demand shocks

$$D_{1,t} = \mu_d + \sum_{j=0}^{\infty} \phi^j \epsilon_{1,t-j}^{true}.$$

Its demand over the leadtime can be expressed as

$$\sum_{i=1}^{\ell_1+1} D_{1,t+i} = (\ell_1+1)\mu_d + \sum_{i=1}^{\ell_1+1} \sum_{j=0}^{\infty} \phi^j \epsilon_{1,t+i-j}^{true} = (\ell_1+1)\mu_d + \sum_{k=0}^{\infty} \omega_{1,k} \epsilon_{t+\ell_1+1-k}^{true} \tag{49}$$

where

$$\omega_{1,k} = \begin{cases} 1 & : \quad k = 0 \\ \omega_{1,k-1} + \phi^k & : \quad 0 < k \leq \ell_1 \\ \phi\omega_{1,k-1} & : \quad k > \ell_1 \end{cases} \tag{50}$$

Hence the retailer's best linear forecast of demand over the leadtime given information available at t is equal to

$$m_{1,t}^{AR} = E\left[\sum_{i=1}^{\ell_1+1} D_{1,t+i}|\mathcal{M}_t^1\right] = (\ell_1+1)\mu_d + \sum_{k=\ell_1+1}^{\infty} \omega_{1,k}\epsilon_{1,t+\ell_1+1-k}^{true}$$

and its forecast error is

$$\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{AR} = \sum_{k=0}^{\ell_1} \omega_{1,k}\epsilon_{t+\ell_1+1-k}^{true}.$$

The MSFE of the retailer's best linear forecast is then given by

$$\nu_1^{AR} = E\left(\sum_{i=1}^{\ell_1+1} D_{1,t+i} - m_{1,t}^{AR}\right)^2 = \sum_{k=0}^{\ell_1} \omega_{1,k}^2\sigma_{\epsilon_1}^2.$$

□

Lemma 2 *The supplier's demand under the suboptimal forecast has an MA(∞) representation with respect to the retailer's demand shocks of*

$$D_{2,t}^{ES} = \mu_d + \sum_{j=0}^{\infty} \psi_j^{ES}\epsilon_{1,t-j}^{ES} \tag{51}$$

where $\epsilon_{1,t-j}^{ES} = [1+\alpha(1+\ell_1)]\epsilon_{1,t-j}^{true}$ *and*

$$\psi_j^{ES} = \frac{(1-\alpha)^j(\alpha^2+\alpha^2\ell_1)}{(\alpha+\phi-1)(1+\alpha+\alpha\ell_1)} + \frac{\phi^j[\phi(1+\alpha+\alpha\phi)-(1+\alpha\ell_1)]}{(\alpha+\phi-1)(1+\alpha+\alpha\ell_1)}, \ \forall j = 0,1,2,\cdots \tag{52}$$

if $\alpha+\phi \neq 1$ *and* $\phi \neq 0$

$$\psi_j^{ES} = \begin{cases} (1-\alpha)^j - \frac{(1+\alpha\ell_1)(1-\alpha)^{j-1}}{1+\alpha+\alpha\ell_1} + \frac{1+\alpha\ell_1}{(1-\alpha)(1+\alpha+\alpha\ell_1)} & : \quad j = 0 \\ (1-\alpha)^j - \frac{(1+\alpha\ell_1)(1-\alpha)^{j-1}}{1+\alpha+\alpha\ell_1} & : \quad j > 0 \end{cases} \tag{53}$$

if $\alpha+\phi \neq 1$ *and* $\phi = 0$

and

$$\psi_j^{ES} = \phi^j(1+j) - \frac{j(1+\alpha\ell_1)\phi^{j-1}}{1+\alpha+\alpha\ell_1}, \forall j = 0, 1, \cdots \tag{54}$$

if $\alpha+\phi = 1$.

Proof of Lemma 2:

Equation (38) can be expressed as

$$D_{2,t}^{ES} = \mu_d + \frac{(1+\alpha(1+\ell_1) - (1+\alpha\ell_1)B}{[1-(1-\alpha)B](1-\phi B)}\epsilon_{1,t}^{true}.$$

Define $G(B) = \frac{1+\alpha(1+\ell_1)-(1+\alpha\ell_1)B}{[1-(1-\alpha)B](1-\phi B)}$, a rational polynomial with the degree of the numerator less than the degree of the denominator. $G(B)$ has different forms of partial fractional expression depending on the sum of α and ϕ and whether value of ϕ is zero. We consider three cases: (i) $\alpha+\phi \neq 1$ and $\phi \neq 0$, (ii) $\alpha+\phi = 1$ and $\phi = 0$, and (iii) $\alpha+\phi = 1$.

Case I: $\alpha + \phi \neq 1$ **and** $\phi \neq 0$**.**

One can express $G(B)$ as

$$G(B) = [1+\alpha(1+\ell_1) - (1+\alpha\ell_1)B]\left[\frac{\Gamma_1}{1-(1-\alpha)B} + \frac{\Gamma_2}{1-\phi B}\right] \tag{55}$$

where

$$\frac{\Gamma_1}{1-1(1-\alpha)B} + \frac{\Gamma_2}{1-\phi B} = \frac{1}{[1-(1-\alpha)B]\,(1-\phi B)}.$$

Solving the equation, we obtain $\Gamma_1 = \frac{\alpha-1}{\alpha+\phi-1}$ and $\Gamma_2 = \frac{\phi}{\alpha+\phi-1}$. Equation (55) then can be expressed as

$$
\begin{aligned}
G(B) &= [1+\alpha(1+\ell_1)-(1+\alpha\ell_1)B]\left[\frac{\Gamma_1}{1-(1-\alpha)B}+\frac{\Gamma_2}{1-\phi B}\right] \\
&= \sum_{j=0}^{\infty}\left\{[1+\alpha(1+\ell_1)]\Gamma_1(1-\alpha)^j+[1+\alpha(1+\ell_1)]\Gamma_2\phi^j\right\}B^j \\
&- \sum_{j=0}^{\infty}[(1+\alpha\ell_1)\Gamma_1(1-\alpha)^j+(1+\alpha\ell_1)\Gamma_2\phi^j]B^{j+1} \\
&= \sum_{j=0}^{\infty}\left\{[1+\alpha(1+\ell_1)]\Gamma_1(1-\alpha)^j+[1+\alpha(1+\ell_1)]\Gamma_2\phi^j\right\}B^j \\
&- \sum_{j=1}^{\infty}[(1+\alpha\ell_1)\Gamma_1(1-\alpha)^{j-1}+(1+\alpha\ell_1)\Gamma_2\phi^{j-1}]B^j \\
&= \sum_{j=0}^{\infty}\left\{[1+\alpha(1+\ell_1)]\Gamma_1(1-\alpha)^j+[1+\alpha(1+\ell_1)]\Gamma_2\phi^j\right\}B^j \\
&- \sum_{j=0}^{\infty}[(1+\alpha\ell_1)\Gamma_1(1-\alpha)^{j-1}+(1+\alpha\ell_1)\Gamma_2\phi^{j-1}]B^j \\
&+ \frac{(1+\alpha\ell_1)\Gamma_1}{1-\alpha}+\frac{(1+\alpha\ell_1)\Gamma_2}{\phi} \\
&= \sum_{j=0}^{\infty}\left\{[1+\alpha(1+\ell_1)]\Gamma_1(1-\alpha)^j+[1+\alpha(1+\ell_1)]\Gamma_2\phi^j\right. \\
&- \left.(1+\alpha\ell_1)\Gamma_1(1-\alpha)^{j-1}-(1+\alpha\ell_1)\Gamma_2\phi^{j-1}\right\}B^j.
\end{aligned}
$$

Note that $\phi \in (-1,1)$, ϕ may be equal to zero. If $\phi \neq 0$, then $\frac{(1+\alpha\ell_1)\Gamma_1}{1-\alpha}+\frac{(1+\alpha\ell_1)\Gamma_2}{\phi}=0$ since its numerator $(1+\alpha\ell_1)\Gamma_1\phi+(1+\alpha\ell_1)\Gamma_2(1-\alpha)=0$. Therefore, $D_{2,t}^{ES}$ has a MA(∞) representation with respect to $\{\epsilon_{1,t}^{true}\}$

$$
D_{2,t}^{ES}=\mu_d+G(B)\epsilon_{1,t}^{true}=\mu_d+\sum_{j=0}^{\infty}\tilde{\psi}_j^{ES}\epsilon_{1,t-j}^{true}. \tag{56}
$$

where

$$\begin{aligned}\tilde{\psi}_j^{ES} &= [1+\alpha(1+\ell_1)]\Gamma_1(1-\alpha)^j + [1+\alpha(1+\ell_1)]\Gamma_2\phi^j - (1+\alpha\ell_1)\Gamma_1(1-\alpha)^{j-1} - (1+\alpha\ell_1)\Gamma_2\phi^{j-1} \\ &= \frac{(1-\alpha)^j(\alpha^2+\alpha^2\ell_1)}{\alpha+\phi-1} + \frac{\phi^j[\phi(1+\alpha+\alpha\phi)-(1+\alpha\ell_1)]}{\alpha+\phi-1} \\ &\quad \forall j = 0,1,2,\cdots\end{aligned}$$

Case II: $\alpha + \phi \neq 1$ **and** $\phi = 0$**.**

Since $\phi = 0$, one can express $G(B)$ as

$$\begin{aligned}G(B) &= \sum_{j=0}^{\infty}[1+\alpha(1+\ell_1)](1-\alpha)^j B^j - \sum_{j=0}^{\infty}(1+\alpha\ell_1)(1-\alpha)^{j-1}B^j + \frac{1+\alpha\ell_1}{1-\alpha} \\ &= \sum_{j=1}^{\infty}\left\{[1+\alpha(1+\ell_1)](1-\alpha)^j - (1+\alpha\ell_1)(1-\alpha)^{j-1}\right\}B^j + \frac{1+\alpha\ell_1}{1-\alpha}\end{aligned}$$

and the coefficients in the $MA(\infty)$ representation of $D_{2,t}^{ES}$ are given by

$$\tilde{\psi}_j^{ES} = \begin{cases} [1+\alpha(1+\ell_1)](1-\alpha)^j - (1+\alpha\ell_1)(1-\alpha)^{j-1} + \frac{1+\alpha\ell_1}{1-\alpha} & : \quad j = 0 \\ [1+\alpha(1+\ell_1)](1-\alpha)^j - (1+\alpha\ell_1)(1-\alpha)^{j-1} & : \quad j > 0 \end{cases} \tag{57}$$

Case III: $\alpha + \phi = 1$**.**

If $\alpha + \phi = 1$, then (55) can be expressed as

$$\begin{aligned}G(B) &= [1+\alpha(1+\ell_1) - (1+\alpha\ell_1)B]\frac{1}{(1-\phi B)^2} \\ &= [1+\alpha(1+\ell_1) - (1+\alpha\ell_1)B]\sum_{j=0}^{\infty}(j+1)\phi^j B^j \\ &= [1+\alpha(1+\ell_1)]\sum_{j=0}^{\infty}(1+j)\phi^j B^j - \sum_{j=0}^{\infty}(1+\alpha\ell_1)(j+1)\phi^j B^{j+1} \\ &= \sum_{j=0}^{\infty}\left\{[1+\alpha(1+\ell_1)](1+j)\phi^j - (1+\alpha\ell_1)j\phi^{j-1}\right\}B^j\end{aligned}$$

The coefficients in the $MA(\infty)$ representation of $D_{2,t}^{ES}$ are given by

$$\tilde{\psi}_j^{ES} = \phi^{j-1}\left\{[1+\alpha(1+\ell_1)]\phi(1+j) - j(1+\alpha\ell_1)\right\}, \forall j = 0,1,\cdots \tag{58}$$

When $j = 0$, the three representations of $G(B)$ all have $\tilde{\psi}_0^{ES} = 1+\alpha+\alpha\ell_1$. Let $\epsilon_{1,t-j}^{ES} = \tilde{\psi}_0^{ES}\epsilon_{1,t-j}^{true} = [1+\alpha(1+\ell_1)]\epsilon_{1,t-j}^{true}$ and $\psi_j^{ES} = \frac{\tilde{\psi}_j^{ES}}{\tilde{\psi}_0^{ES}}$.

Then (56) can be represented as

$$D_{2,t}^{ES} = \mu_d + \sum_{j=0}^{\infty}\left(\frac{\tilde{\psi}_j^{ES}}{\tilde{\psi}_0^{ES}}\right)\tilde{\psi}_0^{ES}\epsilon_{1,t-j}^{true} = \mu_d + \sum_{j=0}^{\infty}\psi_j^{ES}\epsilon_{1,t-j}^{ES}. \tag{59}$$

□

Proof of Proposition 10:

Following Equation (51) in Lemma 2, the supplier's demand over the leadtime $\ell_2 + 1$ is

$$\begin{aligned}\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{ES} &= (1+\ell_2)\mu_d + \sum_{i=1}^{\ell_2+1}\sum_{j=0}^{\infty}\psi_j^{ES}\epsilon_{1,t+i-j}^{ES} \\ &= (1+\ell_2)\mu_d + \sum_{k=0}^{\infty}\xi_k^{ES}\epsilon_{t+\ell_2+1-k}^{ES}\end{aligned}$$

where

$$\xi_k^{ES} = \begin{cases} 1 & : \quad k = 0 \\ \xi_{k-1}^{ES} + \psi_k^{ES} & : \quad 0 < k \le \ell_2 \\ \xi_{k-1}^{ES} + \psi_k^{ES} - \psi_{k-\ell_2-1}^{ES} & : \quad k > \ell_2 \end{cases} \tag{60}$$

Since the shocks $\left\{\epsilon_{1,t+l_2+1}^{ES}, \epsilon_{1,t+l_2}^{ES}, \cdots, \epsilon_{1,t+1}^{ES}\right\}$ are not predictable at time t, their conditional expectations are zero. Therefore, the supplier's best linear forecast of future demand over the leadtime is equal to

$$m_{2,t}^{ES} = E\left[\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{ES}|\mathcal{M}_t^2\right] = (1+\ell_2)\mu_d + \sum_{k=\ell_2+1}^{\infty}\xi_k^{ES}\epsilon_{1,t+\ell_2+1-k}^{ES}.$$

The supplier's MSFE is

$$\nu_2^{ES} = \mathrm{Var}\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{ES} - m_{2,t}^{ES}\right) = [1+\alpha(1+\ell_1)]^2\sum_{k=0}^{\ell_2}(\xi_k^{ES})^2\sigma_{\epsilon_1}^2.$$

Proof of Proposition 11:

If the root of the polynomial $\tilde{\theta}(z) = 1 - \left[\frac{\phi(1-\phi^{\ell_1+1})}{1-\phi^{\ell_1+2}}\right] z$ is inside the unit circle, i.e. $\left|\frac{1-\phi^{\ell_1+2}}{\phi(1-\phi^{\ell_1+1})}\right| < 1$, then the supplier's demand $\{D_{2,t}^{AR}\}$ is not invertible with respect to the retailer's shocks $\{\epsilon_{1,t}^{true}\}$. The inequality implies

$$-1 < \frac{1-\phi^{\ell_1+2}}{\phi(1-\phi^{\ell_1+1})} < 1. \tag{61}$$

When $0 < \phi < 1$, the righthand side of (61) is not satisfied. When $-1 < \phi < 0$, to satisfy both sides of (61), ϕ must be less than one and satisfy

$$2\phi^{\ell_1+2} - \phi > 1. \tag{62}$$

If ℓ_1 is odd, we have

$$2\phi^{\ell_1+2} - \phi = |\phi| - 2|\phi^{\ell_1+2}| < |\phi| < 1$$

which conflicts with (62) and thus the supplier's demand is invertible with respect to the retailer's demand shocks $\{\epsilon_{1,t}^{true}\}$. This establishes the proof of (i).

Next we show the root of the polynomial

$$f(\phi) = 2\phi^{\ell_1+2} - \phi - 1$$

is a function of ℓ_1 when ℓ_1 is even. Lemma 3 shows $f(\phi)$ is a square-free polynomial. Hence we can apply Sturm's theorem to identify the intervals where the roots of $f(\phi)$ are located. From Sturm's theorem, the number of sign changes in the Sturm chain at $\phi = -1$ and $\phi = 0$ are two and one respectively. Hence there is exactly one root between -1 and 0. Let $\phi = \kappa(\ell_1)$ be the root of $f(\phi)$, where $\kappa(\ell_1) \in (-1, 0)$ and ℓ_1 is even. Since $f(0) < 0$ and $f(-1) > 0$ when ℓ_1 is even, we infer that for $\phi \in (-1, \kappa(\ell_1))$, $f(\phi) > 0$, and $\phi \in (\kappa(\ell_1), 0)$, $f(\phi) < 0$. This establishes the proof of (ii). □

Lemma 3 *The polynomial $f(x) = 2x^{\ell_1+2} - x - 1$ is a square-free polynomial.*

Proof of Lemma 3:

Assume $x = a$ is a root of $f(x)$. We can write $f(x) = (x - a)q(x)$. If we take the first derivative of

$f(x)$, then $f'(x) = q(x) + (x-a)q'(x)$. If $f(x)$ has a repeating root such as $x = a$, then $q(x)$ can be factorized as $q(x) = (x-a)\gamma(x)$. This implies that $f'(x)$ shares the common factor $(x-a)$ with $f(x)$. Therefore, a sufficient condition for $f(x)$ to be a square free polynomial is that the greatest common divisor (GCD) of $f(x)$ and $f'(x)$ is a constant. One can apply the Euclidean Algorithm to find the GCD of $f(x)$ and $f'(x)$. Using the polynomial long division, we can express

$$f(x) = f'(x)\frac{x}{\ell_1+2} - \left(\frac{\ell_1+1}{\ell_1+2}\right)x - 1.$$

Repeating the polynomial long division for $f'(x)$ with respect to $-\left(\frac{\ell_1+1}{\ell_1+2}\right)x - 1$, we can express

$$f'(x) = \left[-\left(\frac{\ell_1+1}{\ell_1+2}\right)x - 1\right]p(x) + C(\ell_1)$$

where $p(x)$ is a polynomial whose degree is less than the degree of $f'(x)$ and $C(\ell_1)$ is a constant with value depending on ℓ_1. Applying the polynomial long division again for $\left[-\left(\frac{\ell_1+1}{\ell_1+2}\right)x - 1\right]$ with respect to $C(\ell_1)$, we have the remainder equal to zero. Thus we conclude that the GCD of $f(x)$ and $f'(x)$ is the constant $C(\ell_1)$. □

Lemma 4 *The supplier's demand process under the optimal forecast has the MA(∞) representation with respect to* $\{\epsilon^{true}_{1,t-j}\}$

$$D^{AR}_{2,t} = \mu_d + \sum_{j=0}^{\infty} \psi^{AR}_j \epsilon^{AR,S}_{1,t-j},$$

where

$$\psi^{AR}_j = \begin{cases} 1 & : \quad j = 0 \\ \frac{(1-\phi)\phi^{\ell_1+j+1}}{1-\phi^{\ell_1+2}} & : \quad j \geq 1 \end{cases} \tag{63}$$

Proof of Lemma 4:

Using Equation (46), one can express the supplier's demand under the optimal forecast as

$$
\begin{aligned}
D_{2,t}^{AR} &= \mu_d + \left(1 + \phi \frac{1-\phi^{\ell_1+1}}{1-\phi}\right) \epsilon_{1,t}^{true} + \sum_{j=0}^{\infty} \phi^{\ell_1+2+j} \epsilon_{1,t-1-j}^{true} \\
&= \mu_d + \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right) \epsilon_{1,t}^{true} + \sum_{j=0}^{\infty} \phi^{\ell_1+2+j} \epsilon_{1,t-1-j}^{true} \\
&= \mu_d + \sum_{j=0}^{\infty} \tilde{\psi}_j^{AR} \epsilon_{1,t-j}^{true}
\end{aligned}
$$

where

$$
\tilde{\psi}_j^{AR} = \begin{cases} \left(1 + \sum_{k=1}^{\ell_1+1} \phi^k\right) & : \quad j = 0 \\ \phi^{\ell_1+1+j} & : \quad j \geq 1 \end{cases}
$$

Let $\psi_j^{AR} = \frac{\tilde{\psi}_j^{AR}}{\tilde{\psi}_0^{AR}}$ and $\epsilon_{1,t-j}^{AR,S} = \tilde{\psi}_0^{AR} \epsilon_{1,t-j}^{true}$. Then

$$
D_{2,t}^{AR} = \mu_d + \sum_{j=0}^{\infty} \psi_j^{AR} \epsilon_{1,t-j}^{AR,S}
$$

where

$$
\psi_j^{AR} = \begin{cases} 1 & : \quad j = 0 \\ \frac{(1-\phi)\phi^{\ell_1+j+1}}{1-\phi^{\ell_1+2}} & : \quad j \geq 1 \end{cases}
$$

□

Proof of Proposition 12:

Following Lemma 4, the supplier's demand over the leadtime periods $\ell_2 + 1$ can be expressed as

$$
\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} = (\ell_2+1)\mu_d + \sum_{i=1}^{\ell_2+1} \sum_{j=0}^{\infty} \psi_j^{AR} \epsilon_{1,t+i-j}^{true} = (\ell_2+1)\mu_d + \sum_{k=0}^{\infty} \xi_k^{AR} \epsilon_{1,t+\ell_2+1-k}^{true}
$$

where

$$
\xi_k^{AR} = \begin{cases} 1 & : \\ \xi_{k-1}^{AR} + \psi_k^{AR} & : \quad 0 < k \leq \ell_2 \\ \xi_{k-1}^{AR} + \psi_k^{AR} - \psi_{k-\ell_2-1}^{AR} & : \quad k > \ell_2 \end{cases} \tag{64}
$$

Hence the supplier's best linear forecast of future demand over leadtime is equal to

$$m_{2,t}^{AR,S} = E\left[\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR}|\mathcal{M}_t^2\right] = (1+\ell_2)\mu_d + \sum_{k=\ell_2+1}^{\infty} \xi_k^{AR}\epsilon_{1,t+\ell_2+1-k}^{true}$$

and the supplier's MSFE is

$$\nu_2^{AR,S} = E\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} - m_{2,t}^{AR,S}\right)^2 = \sum_{k=0}^{\ell_2}(\xi_k^{AR})^2\sigma_{\epsilon_1}^2$$

□

Lemma 5 *If the retailer uses optimal forecasting and the root of the polynomial* $\tilde{\theta}(z) = 1 - \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell_1+2}}\right)z$ *is inside the unit circle and the retailer does not share its demand shocks with the supplier, the supplier's demand is equal to*

$$D_{2,t}^{AR} = \mu_d + \sum_{j=0}^{\infty}\psi_{2,j}\tilde{\epsilon}_{2,t-j} \tag{65}$$

where

$$\psi_{2,j} = \begin{cases} 1 & : \quad j=0 \\ \frac{\phi^{j-2}(1-\phi)\left(\phi^{\ell_1+2}-\phi-1\right)}{1-\phi^{\ell_1+1}} & : \quad j \geq 1 \end{cases} \tag{66}$$

and

$$\tilde{\epsilon}_{2,t-j} = \left[\frac{1}{1-\left(\frac{1-\phi^{\ell_1+2}}{\phi-\phi^{\ell_1+2}}\right)B}\right]\left[(1-\phi B)D_{2,t}^{AR} - d\right].$$

Proof of Lemma 5:

When the retailer adopts the optimal forecast, the supplier faces ARMA(1,1) demand specified by Equation (47)

$$\begin{aligned}(1-\phi B)D_{2,t}^{AR} &= d + \left[1 + \sum_{j=1}^{\ell_1+1}\phi^j(1-B)\right]\epsilon_{1,t}^{true} \\ &= d + \left(1+\sum_{j=1}^{\ell_1+1}\phi^j\right)\left[1 - \left(\frac{\sum_{j=1}^{\ell_1+1}\phi^j}{1+\sum_{j=1}^{\ell_1+1}\phi^j}\right)B\right]\epsilon_{1,t}^{true}. \end{aligned} \tag{67}$$

Since the retailer's shocks are not invertible with respect to $\{\epsilon_{1,t}^{true}\}$ and the retailer does not share $\{\epsilon_{1,t}^{true}\}$ with the supplier, the supplier will construct a new ARMA(1,1) representation for its demand with respect to a set of shocks which generates the same linear past as its observed demand (see Brockwell and Davis 1991, pp.125 - 126). Thus Equation (67) becomes

$$\begin{aligned}(1-\phi B)D_{2,t}^{AR} &= d + \left(1+\sum_{j=1}^{\ell_1+1}\phi^j\right)\left[1-\left(\frac{1+\sum_{j=1}^{\ell_1+1}\phi^j}{\sum_{j=1}^{\ell_1+1}\phi^j}\right)B\right]\epsilon_{2,t} \\ &= d + \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)\left[1-\left(\frac{1-\phi^{\ell_1+2}}{\phi-\phi^{\ell_1+2}}\right)B\right]\epsilon_{2,t} \qquad (68)\end{aligned}$$

where

$$\epsilon_{2,t} = \left(\frac{1-\phi}{1-\phi^{\ell_1+2}}\right)\left[\frac{1}{1-\left(\frac{1-\phi^{\ell_1+2}}{\phi-\phi^{\ell_1+2}}\right)B}\right]\left[(1-\phi B)D_{2,t}^{AR}-d\right].$$

Hence the supplier's demand from Equation (68) can be expressed as

$$\begin{aligned}D_{2,t}^{AR} &= \mu_d + \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)\epsilon_{2,t} + \sum_{j=1}^{\infty}\left[\left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)\phi^j - \left[\frac{(1-\phi^{\ell_1+2})^2}{(1-\phi)(\phi-\phi^{\ell_1+2})}\right]\phi^{j-1}\right]\epsilon_{2,t-j} \\ &= \mu_d + \sum_{j=0}^{\infty}\psi_{2,j}\tilde{\epsilon}_{2,t-j}\end{aligned}$$

where

$$\psi_{2,j} = \begin{cases} 1 & : \quad j=0 \\ \frac{\phi^{j-2}(1-\phi)\left(\phi^{\ell_1+2}-\phi-1\right)}{1-\phi^{\ell_1+1}} & : \quad j\geq 1 \end{cases} \qquad (69)$$

and $\tilde{\epsilon}_{2,t-j} = \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)\epsilon_{2,t}$. □

Proof of Proposition 13:

Using Equation (65) from Lemma 5, the Supplier's demand over the leadtime can be expressed as

$$\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} = (\ell_2+1)\mu_d + \sum_{i=1}^{\ell_2+1}\sum_{j=0}^{\infty}\psi_{2,j}\tilde{\epsilon}_{2,t+i-j} = (\ell_2+1)\mu_d + \sum_{k=0}^{\infty}\xi_{2,k}\tilde{\epsilon}_{2,t+\ell_2+1-k}$$

where

$$
\xi_{2,k} = \begin{cases} 1 & : \quad k = 0 \\ \xi_{2,k-1} + \psi_{2,k} & : \quad 0 < k \le \ell_2 \\ \xi_{k-1} + \psi_{2,k} - \psi_{2,k-\ell_2-1} & : \quad k > \ell_2 \end{cases} \tag{70}
$$

The supplier's best linear forecast of its demand over the leadtime $\ell_2 + 1$ is

$$
m_{2,t}^{AR,NS} = (1+\ell_2)\mu_d + \sum_{k=\ell_2+1}^{\infty} \xi_{2,k}\tilde{\epsilon}_{2,t+\ell_2+1-k}.
$$

It is straightforward to show that the supplier's MSFE is equal to

$$
\nu_2^{AR,NS} = E\left(\sum_{i=1}^{\ell_2+1} D_{2,t+i}^{AR} - m_{2,t}^{AR,NS}\right)^2 = \left(\frac{1-\phi^{\ell_1+2}}{1-\phi}\right)^2 \sum_{k=0}^{\ell_2} (\xi_{2,k})^2 \sigma_{\epsilon_2}^2 = \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi}\right)^2 \sum_{k=0}^{\ell_2} (\xi_{2,k})^2 \sigma_{\epsilon_1}^2 \tag{71}
$$

where $\sigma_{\epsilon_2}^2$ is the variance of the supplier's demand shocks $\{\epsilon_{2,t}\}$ and $\sigma_{\epsilon_2}^2 = \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell_1+2}}\right)^2 \sigma_{\epsilon_1}^2$. □

Remark 4 *The condition under which the root is inside the unit circle,* $\left|\frac{1-\phi^{\ell_1+2}}{\phi-\phi^{\ell_1+2}}\right| < 1$, *implies* $\sigma_{\epsilon_2}^2 = \left(\frac{\phi-\phi^{\ell_1+2}}{1-\phi^{\ell_1+2}}\right)^2 \sigma_{\epsilon_1}^2 > \sigma_{\epsilon_1}^2$.

References

[1] P. Brockwell, R.Davis. 1991. Time Series: Theory and Methods, second edition, *Springer-Verlag New York, Inc.*

[2] V. Gaur, A. Giloni, S. Seshadri. 2005. Information sharing in a supply chain under ARMA demand. *Management Science* **5**1, No. 6, 961-969.

[3] A. Giloni, C. Hurvich, S. Seshadri. 2014. Forecasting and information sharing in supply chains under ARMA demand. *IIE Transactions* **4**6, 35-54.

[4] H.L. Lee, K.T. So, C.S. Tang. 2000. The value of information sharing in a two-level supply chain. *Management Science* **4**6, 626-643.

[5] S. Nahmias, T.L. Olsen. 2015. Production and Operations Analysis, seventh edition, *Waveland Press, Inc.*

[6] S. Raghunathan. 2001. Information Sharing in a supply chain: a note on its value when demand is nonstationary. *Management Science* **4**7, 605-610.

[7] M. Shniderman and F. El Ouardighi 2014. The impact of partial information sharing in a two-echelon supply chain. *Operations Research Letters* **4**2 No. 3, 234-237.

[8] E. Silver and R. Petersen. 1985. Decision Systems for Inventory Management and Production Planning, second edition, *John Wiley & Sons, New York, NY.*

[9] Zhang X. 2004. Evoluation of ARMA demand in supply chains. *Manufacturing Service Operation Management* **6**, 195-198.

[10] J.Thomas. 1941. Sturm's Theorem for Multiple Roots. *National Mathematics Magazine* **1**5, No.8, 391-394.