Is there a benefit to sharing market sales information? Linking theory and practice

Takamichi Hosoda*   Mohamed M. Naim   Stephen M. Disney   Andrew Potter

Logistics Systems Dynamics Group,Cardiff Business School,
Cardiff University, Cardiff, Wales, UK

Abstract

Using a real-life data set, we investigate the benefit of sharing market sales information in a setting where a theoretical model argues there is no benefit from such a collaboration scheme. The set of real Electronic Point Of Sales (EPOS) data and the orders that were placed by a retailer to a suppler was used. We have focused on products that operate under an every day low price strategy. To measure the benefit of the second echelon player the Standard Deviation of the Prediction Errors (SDPE) is used as this is linearly related to inventory costs. It is revealed that the second echelon player can reduce its SDPE by between $8 - 19\%$ by exploiting the shared EPOS data, suggesting that there is a benefit to information sharing. Furthermore, it is proposed that the noise element that is originally contained in the EPOS series is the major source of the information sharing benefit.

---

*Corresponding author. *E-mail address:* hosodat@cardiff.ac.uk

Is there a benefit to sharing market sales information? Linking theory and practice

July 27, 2007

Abstract

Using a real-life data set, we investigate the benefit of sharing market sales information in a setting where a theoretical model argues there is no benefit from such a collaboration scheme. The set of real Electronic Point Of Sales (EPOS) data and the orders that were placed by a retailer to a suppler was used. We have focused on products that operate under an every day low price strategy. To measure the benefit of the second echelon player the Standard Deviation of the Prediction Errors (SDPE) is used as this is linearly related to inventory costs. It is revealed that the second echelon player can reduce its SDPE by between $8 - 19\%$ by exploiting the shared EPOS data, suggesting that there is a benefit to information sharing. Furthermore, it is proposed that the noise element that is originally contained in the EPOS series is the major source of the information sharing benefit.

1 Introduction

"The benefit of sharing market sales information in a supply chain" has received a lot of research attention recently via a variety of approaches. Among them, time-series analysis is a popular research tool. Graves (1999), Gavirneni *et al.* (1999), Lee *et al.* (2000), Raghunathan (2001), Aviv (2002), Alwan *et al.* (2003), Zhang (2004), Gilbert (2005), Gaur *et al.* (2005), and Hosoda and Disney (2006) are typical examples that exploit this technique. All these papers have used a serially linked multi-echelon supply chain model. The common assumptions among those papers are that:

- The external market sales can be represented as an ARIMA (Box *et al.*, 1994) process,

- A serially linked supply chain, periodic review replenishment policy with a constant lead-time system exists,

- Unmet demand at the first echelon (the retailer) is backlogged, and the upper echelon player (the supplier) can meet all demand,

- All players in a system exploit the Order-Up-To (OUT) policy,

- Minimum Mean Square Error (MMSE) forecasting is used within the OUT policy to generate inventory position targets.

It is well recognised that under the assumptions stated above, the system is linear and thus, the process of the order placed by the first echelon player is mathematically tractable. For example, if the sales pattern follows an AR(1) structure, then the order process generated through the OUT policy is an ARMA(1, 1) process. Furthermore, this ARMA(1, 1) process includes all available information on the market sales process. Therefore, if the second echelon player exploits an appropriate filter, theoretically, this player can extract all the necessary information about the market sales without observing it directly. Based on this reasoning, several researchers have concluded that the benefit of the market sales information sharing is at best minor (see, Graves, 1999; Raghunathan, 2001; Alwan *et al.*, 2003; Zhang, 2004; Gilbert, 2005; Hosoda and Disney, 2006, for example). From a practical point of view, however, a natural question arises. Under a real setting, is it still possible for the second echelon player to extract the information of the market sales process with a usable level of accuracy? In other words, we would like to know whether the benefit of the market sales information sharing exists in a real supply chain. To measure the benefit of the sales information sharing, we will use the Standard Deviation of the Prediction Errors (SDPE). This measure is a good indicator of the inventory related costs when inventory holding and backlog costs are linear in the stock position. This occurs when the level of the safety stock has been set via the newsboy approach.

This research paper is organised as follows. First the data sets collected from a retail supply chain will be described. Then a model to analyse the benefit of the EPOS data sharing will be provided and the results of the time series analysis will be shown. In the time series analysis, we will first show estimated structures of the EPOS process. Next, we will identify the theoretical structure of the ORDER series. This theoretical structure will be applied into the real set of the ORDER data to filter out noise and to estimate the values of the essential parameters. We quantify the benefit of the information sharing using the SDPE measure. Finally, we will conclude our analysis.

2   A case from a retail supply chain

Data has been collected from a soft drink manufacturer, called DrinkCo for anonymity, through a comprehensive diagnostic activity termed the Quick Scan[1]. DrinkCo has been involved in a Co-Managed Inventory (CMI) scheme (Christopher, 1998) with a major UK grocery retailer, termed

---

[1]The Quick Scan will not be reviewed here. Instead, we refer the reader to Naim *et al.* (2002).

SuperStore for anonymity. SuperStore is the largest customer of DrinkCo and one of the largest retailers in the UK. Through the Quick Scan activity, the historic Electronic Point Of Sales (EPOS) data and the historic order data placed by SuperStore to DrinkCo, termed the ORDER, have been obtained. The ordering policy used by SuperStore has been investigated by Potter *et al.* (2004) and it is the OUT policy. To generate forecasts, SuperStore exploits a state-of-the-art information system. This system seeks to minimise forecast errors. Therefore, we have employed an MMSE forecasting technique herein to represent the retailers forecasting method, as MMSE forecasts are mathematically tractable. It should be noted, however, that in practise, the market demand forecasting draws on information regarding not only the EPOS data but also from market intelligence (Potter *et al.*, 2005). For example, if there is a period that sales of soft drinks will be expected to increase significantly (for example, periods of hot weather or public holidays) then to maintain the customer service levels, the original forecast will be modified. The replenishment lead-times in the supply chain under investigation are unity (and this includes the review period). DrinkCo is producing approximately 200 different product lines and among them, three products with a low, medium and high volume of sales were selected. In this paper, the three products selected are called Low, Med. and High, respectively. Each product is carrying the SuperStore's brand name and is subject to an Every Day Low Price (EDLP) strategy. As such, there are no sales promotions for these products. Fig. 1 shows the time-series plots of the EPOS and the ORDER data utilised in the analysis. All data is aggregated by week, starts in the middle of May 2003 and is 51 weeks long. Table 1 presents some descriptive statistics. In both Fig. 1 and Table 1, the bullwhip phenomenon (see, Geary *et al.*, 2006, for example) is observed. In addition, during the summer period a high volume of sales is observed due to a very hot summer. Peak temperatures occurred during weeks of 10 to 15 and this is reflected in the EPOS data with high demand peaks for all three products. Furthermore, a large disturbance in the ORDER series can be observed during the Christmas season (week 31–35).
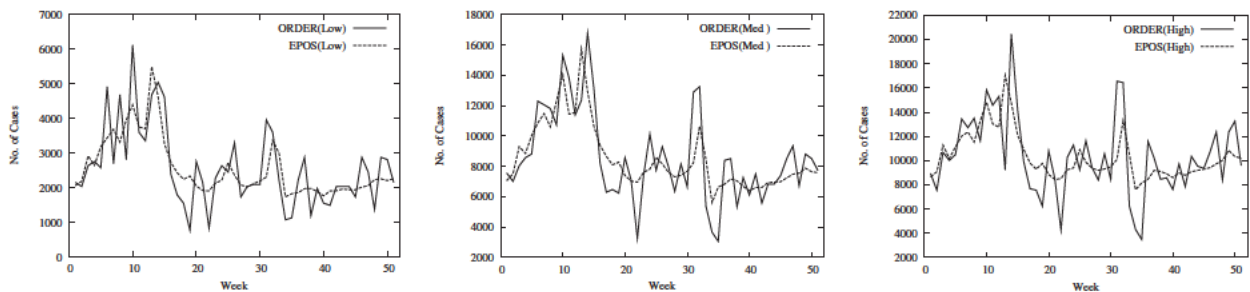

Figure 1: Time series plots of EPOS and ORDER

Table 1: Descriptive statistics of EPOS and ORDER

		Weekly				Standard	Coefficient
		Average	Min.	Max.	Variance	deviation	of variation
Low	EPOS	2,563	1,742	5,493	692,014	832	0.325
	ORDER	2,578	780	6,120	1,300,831	1,141	0.442
Med.	EPOS	8,547	5,612	15,764	4,566,748	2,137	0.250
	ORDER	8,578	3,060	16,800	8,704,675	2,950	0.344
High	EPOS	10,320	7,595	17,048	3,682,693	1,919	0.186
	ORDER	10,390	3,480	20,460	10,668,806	3,266	0.314

3 The model

In this paper, we will exploit the following notation;

X_t : EPOS rate realised at time t.
Y_t : ORDER rate realised at time t.
d : Mean of the EPOS process.
ρ : Autoregressive (AR) parameter.
θ : Moving Average (MA) parameter.
ε_t, ν_t : Noise term at time t in an AR(1) and ARMA(1, 1) processes, respectively.
σ_ε : Standard deviation of ε_t.
$\hat{X}_t$: MMSE forecast of X_t.
$\hat{Y}_t$: MMSE forecast of Y_t.
$\hat{Y}_{t,\mathbf{E}}$: MMSE forecast of Y_t induced by the EPOS data.
$\hat{Y}_{t,\mathbf{O}}$: MMSE forecast of Y_t induced by the ORDER data.
S_t : Order-Up-To (OUT) level at time period t.
$\hat{d}_{\mathbf{E}}$: Estimated value of d using the EPOS data.
$\hat{d}_{\mathbf{O}}$: Estimated value of d using the ORDER data.
$\hat{\rho}_{\mathbf{E}}$: Estimated value of ρ using the EPOS data.
$\hat{\rho}_{\mathbf{O}}$: Estimated value of ρ using the ORDER data.
$\hat{\theta}_{\mathbf{O}}$: Estimated value of θ using the ORDER data.
$\hat{\varepsilon}_t$: Estimated value of ε_t using the ORDER data.
n : Length of the time series data.
ω_t : Effect of intervention event at time t.
$P_t^{(T)}$: Pulse function $= \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases}$, where T represents the intervention period.
SDPE : Standard Deviation of the Prediction Errors $= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}$.

Let us begin with the theory that results from the assumption of the EPOS series. If the EPOS follows an AR(1) process

$$X_{t+1} = d + \rho(X_t - d) + \varepsilon_{t+1}, \tag{1}$$

the MMSE forecast of the lead-time sales is given by a conditional expectation and is

$$\hat{X}_{t+1} = E[X_{t+1}|d, \rho, X_t] = d + \rho(X_t - d).$$

Using time variant OUT levels at time period t (S_t), the traditional OUT policy can be written as

$$Y_t = X_t + (S_t - S_{t-1}), \tag{2}$$

where $S_t = \hat{X}_{t+1} +$ *safety stock*. Assuming SuperStore uses MMSE forecasting, the ORDER series becomes an ARMA(1, 1) process

$$Y_{t+1} \quad = \quad d + \rho(Y_t - d) + (1 + \rho)\varepsilon_{t+1} - \rho\varepsilon_t, \tag{3}$$

as is shown in Appendix A. Comparing Eq. 1 and Eq. 3 reveals that the values of the parameters in the EPOS process, such as ρ, d and ε_t, are contained in the ORDER process. Therefore, in theory, if DrinkCo knows that the EPOS follows an AR(1) process and SuperStore uses the OUT policy with an MMSE forecasting method, then DrinkCo can create an MMSE forecast of the ORDER by exploiting an ARMA(1, 1) model structure *without* sharing the up-to-date EPOS data. From Eq. 3, the MMSE forecast of Y_{t+1} can be written as

$$\hat{Y}_{t+1} = E[Y_{t+1}|d, \rho, \varepsilon_t, Y_t] = d + \rho(Y_t - d) - \rho\varepsilon_t. \tag{4}$$

The difference between Eq. 3 and Eq. 4, $Y_{t+1} - \hat{Y}_{t+1}$, represents the forecast error and in this case it is $(1 + \rho)\varepsilon_{t+1}$. It should be also noted that Eq. 3 can be rewritten as

$$Y_{t+1} = d + \rho^2(X_t - d) + (1 + \rho)\varepsilon_{t+1}, \tag{5}$$

as shown in Appendix A, so that when X_t is available to DrinkCo directly, the MMSE forecast of Y_{t+1} is then given by

$$\hat{Y}_{t+1} = E[Y_{t+1}|d, \rho, X_t] = d + \rho^2(X_t - d). \tag{6}$$

Both Eq. 4 (no information sharing case) and Eq. 6 (information sharing case) are induced from the common assumption that the market sales follows an AR(1) process. When the market sales follows another type of ARIMA process, different expressions should be used to obtain MMSE forecasts, and these expressions are easily obtained through the same steps as shown in Appendix A. In this sense, the assumption of an AR(1) market sales process is not severe.

From Eq. 6, we may conclude that DrinkCo can also generate an MMSE forecasting using the shared up-to-date EPOS data (X_t) and the values of ρ and d. Subtracting Eq. 6 from Eq. 5 reveals

that when the EPOS data is shared, the forecast error is $(1 + \rho)\varepsilon_{t+1}$. This is still the same as the no information case that is subtracting Eq. 4 from Eq. 3. Hence, in both information sharing and no information sharing cases, the infinite values of SDPE's are the same: $(1 + \rho)\sigma_\varepsilon$. From practical point of view, however, it would be more reasonable to expect that the value of SDPE given by Eq. 4 and that by Eq. 6 may not be the same. Thus we will calculate and compare these values to measure the benefit of the sales information sharing. If DrinkCo has no up-to-date market sales information but knows that the structure of the market sales is AR(1), DrinkCo may use the following expression to conduct an MMSE forecast of Y_{t+1};

$$\hat{Y}_{t+1,\mathbf{O}} = E[Y_{t+1}|\hat{d}_\mathbf{O}, \hat{\rho}_\mathbf{O}, \hat{\varepsilon}_t, Y_t] = \hat{d}_\mathbf{O} + \hat{\rho}_\mathbf{O}(Y_t - \hat{d}_\mathbf{O}) - \hat{\rho}_\mathbf{O}\hat{\varepsilon}_t, \tag{7}$$

where $\hat{\varepsilon}_t = (Y_t - \hat{Y}_{t,\mathbf{O}})/(1 + \hat{\rho}_\mathbf{O})$, and $\hat{Y}_{t+1,\mathbf{O}}$ represents an MMSE forecast of Y_{t+1} without using the EPOS data (X_t). On the other hand, if the updated value of X_t is available to DrinkCo and $\hat{d}_\mathbf{E}$ and $\hat{\rho}_\mathbf{E}$ are also provided by SuperStore [2], DrinkCo can use the following expression;

$$\hat{Y}_{t+1,\mathbf{E}} = E[Y_{t+1}|\hat{d}_\mathbf{E}, \hat{\rho}_\mathbf{E}, X_t] = \hat{d}_\mathbf{E} + \hat{\rho}_\mathbf{E}^2(X_t - \hat{d}_\mathbf{E}), \tag{8}$$

where $\hat{Y}_{t+1,\mathbf{E}}$ is an MMSE forecast of Y_{t+1} relying on the EPOS data (X_t). As a measure of performance, the difference ratio of SDPE's is interesting here, given by

$$\Delta\text{SDPE} = \frac{\text{SDPE}_\text{O} - \text{SDPE}_\text{E}}{\text{SDPE}_\text{O}},$$

where SDPE_O and SDPE_E are $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_{i,\mathbf{O}})^2}$ and $\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_{i,\mathbf{E}})^2}$, respectively. The model we have presented is based on the precise model identification of the market sales process. In Appendix B, we discuss the impact of model misidentification on the value of ΔSDPE.

4 Model structure identification

In this section, the results of time series analysis on the EPOS series will first be shown. Then, we will present the theoretically obtained structures of the ORDER series, based on the identified EPOS structures and generate an expression for the MMSE forecasting.

Estimation of the EPOS structures Fig. 2 shows the AutoCorrelation Functions (ACF's) and the Partial ACF's (PACF's) of the EPOS series. The ACF's disqualify a moving average process, and the PACF's suggest an AR(1) structure, though the third lag PACF is almost significant.

---

[2]Alternatively, it is also reasonable to assume that the DrinkCo estimates those two parameters using the shared historical data of X_t.
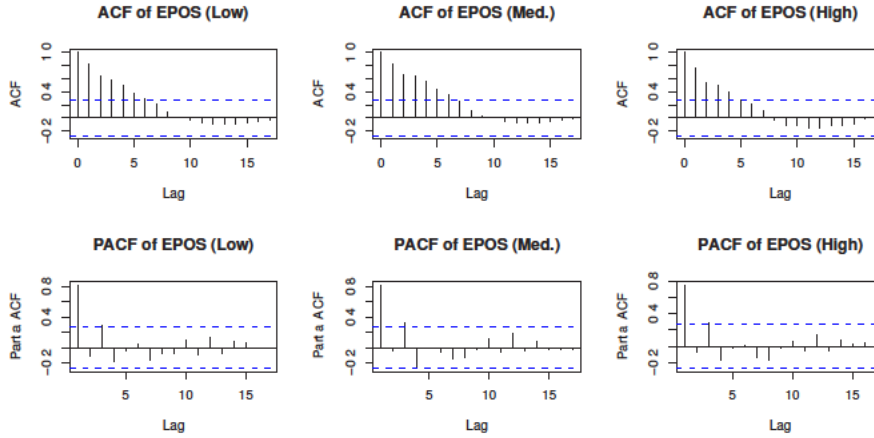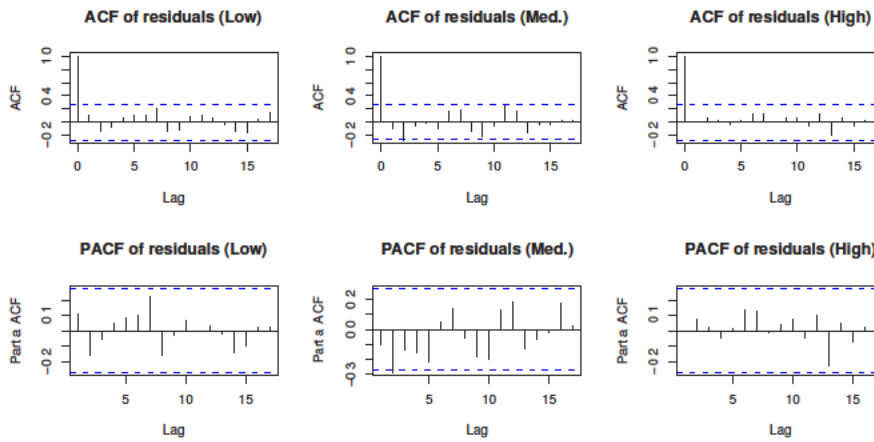
Figure 2: ACF and PACF plot of EPOS's

Figure 3: AR(1) EPOS model diagnostics: ACF and PACF plot of residuals

Three different structures, therefore, are applied; AR(1), AR(2) and AR(3). In the outlier detection procedure, we found a large magnitude of residual at time $t = 13$. Details of these three models are shown in Table 2. The AIC_C (Hurvich and Tsai, 1989) is used as a model selection criterion, since it provides better model choices than other criteria (such as AIC) when the sample size is small (Hurvich and Tsai, 1989). After computing the AIC_C, we selected the AR(1) structures for all three items (Table 2). Diagnostic checks on the residuals revealed no model inadequacy (Fig. 3). Thus, we may reasonably expect that the ORDER series should have an ARMA(1, 1) structure with the parameters of $d = \hat{d}_E$, $\rho = \hat{\rho}_E$ and $\theta = \hat{\rho}_E/(1 + \hat{\rho}_E)$, using the knowledge shown in the previous section.

Applying an ARMA(1, 1) structure to the ORDER data   All the ACF's and the PACF's shown in Fig. 4 are compatible with our expectation that the ORDER series follows an ARMA(1, 1) process. It is recognised that just before and after Christmas (from week 31 to 35) there is a surge

Table 2: Estimated EPOS models

Model structures		Estimated models	ω_{13}	AIC_C
Estimated EPOS(Low) models	AR(1)	$X_{t+1} = 2474^\dagger + 0.839^\dagger(X_t - 2474^\dagger - \omega_{13}P_t^{(13)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$1354^\dagger$	762.260
	AR(2)	$X_{t+1} = 2487^\dagger + 0.938^\dagger(X_t - 2487^\dagger - \omega_{13}P_t^{(13)})$ $-0.113(X_{t-1} - 2487^\dagger - \omega_{13}P_t^{(14)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$1278^\dagger$	764.160
	AR(3)	$X_{t+1} = 2483^\dagger + 0.987^\dagger(X_t - 2483^\dagger - \omega_{13}P_t^{(13)})$ $-0.368(X_{t-1} - 2483^\dagger - \omega_{13}P_t^{(14)})$ $+0.240(X_{t-2} - 2483^\dagger - \omega_{13}P_t^{(15)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$1113^\dagger$	764.096
Estimated EPOS(Med.) models	AR(1)	$X_{t+1} = 8266^\dagger + 0.852^\dagger(X_t - 8266^\dagger - \omega_{13}P_t^{(13)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$3567^\dagger$	855.329
	AR(2)	$X_{t+1} = 8292^\dagger + 0.923^\dagger(X_t - 8292^\dagger - \omega_{13}P_t^{(13)})$ $-0.081(X_{t-1} - 8292^\dagger - \omega_{13}P_t^{(14)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$3466^\dagger$	857.517
	AR(3)	$X_{t+1} = 8271^\dagger + 0.963^\dagger(X_t - 8271^\dagger - \omega_{13}P_t^{(13)})$ $-0.359(X_{t-1} - 8271^\dagger - \omega_{13}P_t^{(14)})$ $+0.274(X_{t-2} - 8271^\dagger - \omega_{13}P_t^{(15)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$2976^\dagger$	856.658
Estimated EPOS(High) models	AR(1)	$X_{t+1} = 10155^\dagger + 0.775^\dagger(X_t - 10155^\dagger - \omega_{13}P_t^{(13)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$3402^\dagger$	863.643
	AR(2)	$X_{t+1} = 10174^\dagger + 0.860^\dagger(X_t - 10174^\dagger - \omega_{13}P_t^{(13)})$ $-0.105(X_{t-1} - 10174^\dagger - \omega_{13}P_t^{(14)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$3268^\dagger$	865.603
	AR(3)	$X_{t+1} = 10190^\dagger + 0.895^\dagger(X_t - 10190^\dagger - \omega_{13}P_t^{(13)})$ $-0.326(X_{t-1} - 10190^\dagger - \omega_{13}P_t^{(14)})$ $+0.238(X_{t-2} - 10190^\dagger - \omega_{13}P_t^{(15)})$ $+\omega_{13}P_t^{(12)} + \varepsilon_{t+1}$	$2919^\dagger$	865.536

$\dagger$: statistically significant at 5% level.

in the ORDER process (see also, Fig. 1). To deal with this, we have incorporated intervention terms into the ORDER models for the Christmas period. In addition, proceeding with the outlier detection procedure, we identify an outlier at time $t = 14$. The estimated parameters are shown in Table 3, with the ACF's and the PACF's of residuals (Fig. 5). The diagnostic check on the residuals supports our ARMA(1, 1) structure assumption. Table 4 summarises the estimation results. It contains two ORDER models; the theoretical model induced from the EPOS analysis and the model obtained through the time series analysis of the ORDER data. The estimated parameters, $\hat{d}_{\mathbf{o}}$ and $\hat{\rho}_{\mathbf{o}}$, are quite close to the values of the theoretically induced parameters. This result supports the fact that AR(1) processes are transformed into ARMA(1, 1) processes, with the predicted values of the parameters. However, the values of $\hat{\theta}_{\mathbf{o}}$ are a little bit different from our expectations. A possible explanation on this phenomenon is discussed in Appendix C.
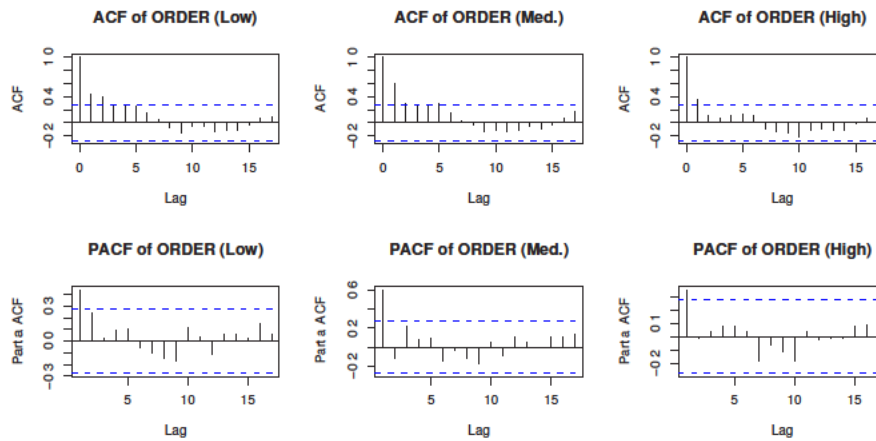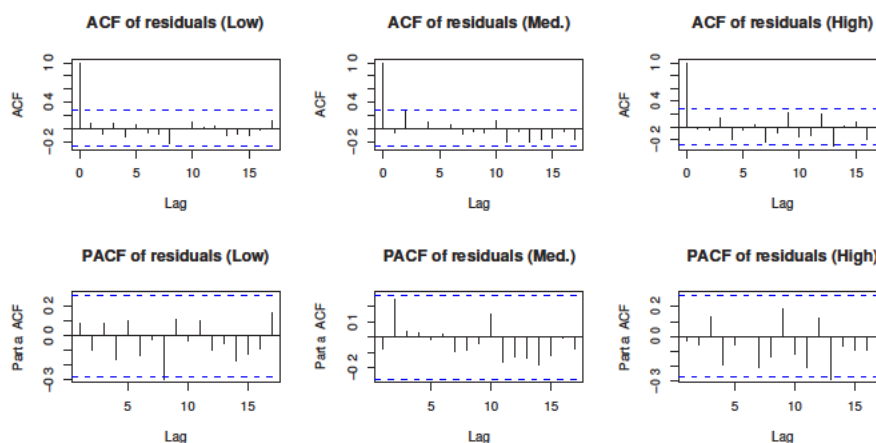

Figure 4: ACF and PACF plot of ORDER's

Figure 5: ARMA(1, 1) ORDER model diagnostics: ACF and PACF plot of residuals

Table 3: Estimated ARMA(1, 1) models for ORDER's

	Estimated models		$\omega_{(\mathbf{Xmas})}$	Outliers
Estimated ORDER (Low) model	$Y_{t+1} =$	$2533^\dagger + 0.836^\dagger \times$ $(Y_t - 2533^\dagger - \omega_{(\mathbf{Xmas})}\mathbf{P_t^{(Xmas)}})$ $+\omega_{(\mathbf{Xmas})}\mathbf{P_t^{(Xmas-1)}} + \nu_{t+1} - 0.510^\dagger\nu_t$	$\omega_{31} : 1746^\dagger$ $\omega_{32} : 1362$ $\omega_{33} : -32$ $\omega_{34} : -1177$ $\omega_{35} : -1112$	
Estimated ORDER (Med.) model	$Y_{t+1} =$	$8386^\dagger + 0.813^\dagger \times$ $(Y_t - 8386^\dagger - \omega_{(\mathbf{Xmas})}\mathbf{P_t^{(Xmas)}} - \omega_{14}P_t^{(14)})$ $+\omega_{(\mathbf{Xmas})}\mathbf{P_t^{(Xmas-1)}} + \omega_{14}P_t^{(13)}$ $+\nu_{t+1} - 0.285\nu_t$	$\omega_{31} : 5522^\dagger$ $\omega_{32} : 5669^\dagger$ $\omega_{33} : -2370$ $\omega_{34} : -4262^\dagger$ $\omega_{35} : -4995^\dagger$	$\omega_{14} : 5047^\dagger$
Estimated ORDER (High) model	$Y_{t+1} =$	$10240^\dagger + 0.757^\dagger \times$ $(Y_t - 10240^\dagger - \omega_{\mathbf{Xmas}}\mathbf{P_t^{(Xmas)}} - \omega_{14}P_t^{(14)})$ $+\omega_{(\mathbf{Xmas})}\mathbf{P_t^{(Xmas-1)}} + \omega_{14}P_t^{(13)}$ $+\nu_{t+1} - 0.434\nu_t$	$\omega_{31} : 6906^\dagger$ $\omega_{32} : 6584^\dagger$ $\omega_{33} : -3788$ $\omega_{34} : -5863^\dagger$ $\omega_{35} : -6854^\dagger$	$\omega_{14} : 8902^\dagger$

$$\omega_{(\mathbf{Xmas})} = [\omega_{31}, \ \omega_{32}, \ \omega_{33}, \ \omega_{34}, \ \omega_{35}]$$
$$\mathbf{P_t^{(Xmas-1)}}' = [P_t^{(30)}, \ P_t^{(31)}, \ P_t^{(32)}, \ P_t^{(33)}, \ P_t^{(34)}]$$
$$\mathbf{P_t^{(Xmas)}}' = [P_t^{(31)}, \ P_t^{(32)}, \ P_t^{(33)}, \ P_t^{(34)}, \ P_t^{(35)}]$$
$$\dagger : \text{ statistically significant at 5\% level.}$$

Table 4: Comparison of theoretically obtained and statistically estimated models

		Theoretically induced ORDER models based on the EPOS data analysis			Estimated ORDER models (ratio to theoretical value)		
Item	Structure	$d(=\hat{d}_\mathbf{E})$	$\rho(=\hat{\rho}_\mathbf{E})$	$\theta(=\hat{\rho}_\mathbf{E}/(1+\hat{\rho}_\mathbf{E}))$	$\hat{d}_\mathbf{O}$	$\hat{\rho}_\mathbf{O}$	$\hat{\theta}_\mathbf{O}$
Low	ARMA(1, 1)	2474	0.839	0.456	2533 (1.024)	0.836 (0.996)	0.510 (1.118)
Med.	ARMA(1, 1)	8266	0.852	0.460	8386 (1.015)	0.813 (0.954)	0.285 (0.620)
High	ARMA(1, 1)	10155	0.775	0.437	10240 (1.008)	0.757 (0.977)	0.434 (0.993)

5   The benefit of information sharing

The calculated SDPE's are shown in Table 5 together with the ΔSDPE's [3]. Since all ΔSDPE's
are positive, we would conclude that there is a benefit in the market sales information sharing in
the SuperStore–DrinkCo supply chain. Let us consider why reality is different from the theory. A
source of the benefit is the forecast accuracy and our analysis shows that the forecast accuracy of
Eq. 8 is much better than that of Eq. 7. To exploit Eq. 7, we need $\hat{\varepsilon}_t$ in addition to $\hat{d}_{\mathbf{o}}$ and $\hat{\rho}_{\mathbf{o}}$.
Since the values of $\hat{d}_{\mathbf{o}}$ and $\hat{\rho}_{\mathbf{o}}$ are quite close to the theoretical values as shown in Table 4, it might
be reasonable to assume that in the case when no information is shared the poor forecast accuracy
is due to a result of a poor estimate of ε_t. On the other hand, Eq. 8 contains ε_t by nature since
X_t can be written as $d + \rho(X_{t-1} - d) + \varepsilon_t$ (see Eq. 1). We might, therefore, conclude that a source
of the benefit of the EPOS data sharing is the fact that the second echelon player can make use
of the true values of the white noise elements, i.e. ε_t, which are originally hidden in the EPOS
process. Without the up-to-date EPOS data, the second echelon player has to estimate the value
of ε_t, which may not be easy to obtain with an acceptable level of accuracy in a practical setting.

Table 5: The values of SDPE's and the benefit of information sharing

	$\mathrm{SDPE_O}$	$\mathrm{SDPE_E}$	ΔSDPE
Low	906.263	737.243	18.65%
Med.	1723.496	1507.767	12.52%
High	2004.169	1843.729	8.01%

6   Conclusion

Using a real-life data set, the benefit of sharing the market sales information has been investigated
in a setting where a theoretical model argues there is no benefit of information sharing. As the
indicator of the benefit, SDPE is exploited. It has been shown that under the case used herein,
sharing EPOS data reduces the second echelon's SDPE by $8 - 19\%$.

A potential source of the benefit has been discussed. We have proposed that the source is the
white noise element that is originally contained in the EPOS series. Without information sharing,
the second echelon player has to estimate the value of that element every period. If the up-to-date
EPOS data (X_t) is shared instead, simply incorporating X_t into the forecast expression enables

---

[3]The estimated intervention variables shown in Table 3 were used when we calculated SDPE's.

the second echelon player to yield more accurate forecasts, as X_t immediately contains the true value of the white noise element. The simple form of the mathematical expression of the MMSE forecast with X_t (Eq. 8) is another advantage of information sharing. This is no more complex than exponential smoothing or moving average forecasting methods.

Acknowledgements

A   The transition of AR(1) process

Since $S_t = \hat{X}_{t+1} + safety\ stock$, the Order-Up-To (OUT) policy (Eq. 2) can be written as

$$
\begin{aligned}
Y_t &= X_t + (S_t - S_{t-1}) \\
&= X_t + (\hat{X}_{t+1} - \hat{X}_t) \\
&= X_t + (d + \rho(X_t - d) - (d + \rho(X_{t-1} - d))) \\
&= X_t + \rho X_t - \rho X_{t-1}.
\end{aligned}
\tag{9}
$$

Incorporating $X_t = d + \rho(X_{t-1} - d) + \varepsilon_t$ and $X_{t-1} = d + \rho(X_{t-2} - d) + \varepsilon_{t-1}$ into Eq. 9 yields

$$
Y_t = d + \rho(X_{t-1} + \rho X_{t-1} - \rho X_{t-2} - d) + (1 + \rho)\varepsilon_t - \rho\varepsilon_{t-1}.
\tag{10}
$$

Since from Eq. 9, we know $X_{t-1} + \rho X_{t-1} - \rho X_{t-2} = Y_{t-1}$, Eq. 10 can be rewritten as

$$
\begin{aligned}
Y_t &= d + \rho(Y_{t-1} - d) + (1 + \rho)\varepsilon_t - \rho\varepsilon_{t-1} \\
&= d + \rho(Y_{t-1} - d) + \nu_t - \theta\nu_{t-1},
\end{aligned}
\tag{11}
$$

where $\nu_t = (1 + \rho)\varepsilon_t$ and $\theta = \rho/(1 + \rho)$. Eq. 11 is an ARMA(1, 1) process. If we incorporate only $X_t = d + \rho(X_{t-1} - d) + \varepsilon_t$ into Eq. 9, we will have another expression of Y_t;

$$
\begin{aligned}
Y_t &= d + \rho(X_{t-1} - d) + \varepsilon_t + \rho(d + \rho(X_{t-1} - d) + \varepsilon_t) - \rho X_{t-1} \\
&= d + \rho^2(X_{t-1} - d) + (1 + \rho)\varepsilon_t,
\end{aligned}
$$

which can be used in the information sharing case.

B   Model structure misidentification and ΔSDPE

Model structure identification of the market sales process is critical in this research. Unfortunately, from a practical point of view, it is reasonable to assume that the perfect identification of the true structure of the market sales is not always achieved, since no models are perfect. In this paper, both Eq. 4 and Eq. 6 are induced by using an assumption that the market sales has an AR(1) structure. If the true structure of the market sales is not AR(1), the forecasts given by these equations are not the MMSE forecasts anymore. Therefore, misidentification will increase the values of SDPE:

$$\text{SDPE}^+ = \text{SDPE} + \gamma,$$

where SDPE^+ is the value of SDPE when the market sales structure is misidentified, and γ is a positive value and is used to represent the magnitude of the impact of model misidentification. We have assumed that the value of γ is identical for both information sharing and no information sharing cases. More detailed analysis about the values of γ will be discussed in our future research. Here, it will suffice to highlight only a simple example that shows that the value of γ is identical in the cases when there is or is no information sharing.

Let's assume that the true structure of the market sales is ARMA(1, 1) and the retailer has misidentified it and believes that it is an AR(1). In addition, this retailer's knowledge about the market sales structure and the value of autoregressive parameter is common knowledge. For the sake of brevity, we will also assume $d = 0$ without loss of generality. The true market sales model can be written as

$$X_t = \rho X_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}, \tag{12}$$

and the model used by the retailer is

$$X_t = \acute{\rho} X_{t-1} + \xi_t,$$

where $\acute{\rho}\ (\neq \rho)$ and ξ_t is the autoregressive coefficient and error term at time t used by the retailer, and from above two expressions, ξ_t can be written as

$$\xi_t = (\rho - \acute{\rho})X_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}. \tag{13}$$

Since the retailer believes that the market sales process is an AR(1), the forecast made by the retailer, $\hat{X}_{t+1}$, can be written as $\acute{\rho} X_t$, which yields the following ARMA(1, 2) ordering process.

$$Y_t = X_t - (S_t - S_{t-1})$$

13

$$= X_t + (\hat{X}_{t+1} - \hat{X}_t)$$

$$= X_t + \acute{\rho} X_t - \acute{\rho} X_{t-1} \tag{14}$$

$$= \rho Y_{t-1} + (1 + \acute{\rho}) \varepsilon_t - (\theta + \acute{\rho}\theta + \acute{\rho}) \varepsilon_{t-1} + \acute{\rho}\theta \varepsilon_{t-2}. \tag{15}$$

To reach this result, we used the same procedure as shown in Appendix A. Since the supplier also believes that the market sales process is an AR(1), he will use an ARMA(1, 1) structure to complete an MMSE forecasting for the order from the retailer. Therefore, the forecast given by the supplier can be written as

$$\hat{Y}_{t+1} = \acute{\rho} Y_t - \acute{\rho} \xi_t, \tag{16}$$

which is for the non-information sharing case and

$$\hat{Y}_{t+1} = \acute{\rho}^2 X_t, \tag{17}$$

for the information sharing case. Both expressions can be obtained using the same method as used for Eq. 4 and Eq. 6, respectively. By taking difference between Eq. 15 and Eq. 16, and after some algebraic simplification with Eq. 13 and Eq. 14, we can have the following expression of the forecast error when no market sale information is shared.

$$
\begin{aligned}
Y_t - \hat{Y}_t &= (\rho Y_{t-1} + (1 + \acute{\rho}) \varepsilon_t - (\theta + \acute{\rho}\theta + \acute{\rho}) \varepsilon_{t-1} + \acute{\rho}\theta \varepsilon_{t-2}) - (\acute{\rho} Y_{t-1} - \acute{\rho} \xi_{t-1}) \\
&= (\rho - \acute{\rho})(1 + \acute{\rho}) X_{t-1} + (1 + \acute{\rho}) \varepsilon_t - \theta (1 + \acute{\rho}) \varepsilon_{t-1}. \tag{18}
\end{aligned}
$$

The forecast error for the information sharing case can be obtained from Eq. 12, Eq. 14 and Eq. 17.

$$
\begin{aligned}
Y_t - \hat{Y}_t &= (X_t + \acute{\rho} X_t - \acute{\rho} X_{t-1}) - (\acute{\rho}^2 X_{t-1}) \\
&= (\rho - \acute{\rho})(1 + \acute{\rho}) X_{t-1} + (1 + \acute{\rho}) \varepsilon_t - \theta (1 + \acute{\rho}) \varepsilon_{t-1},
\end{aligned}
$$

which is identical to the forecast error in non-information sharing case (Eq. 18). Therefore, we may conclude that the impact of misidentification on the forecast errors, γ, is the same for both information and non-information sharing cases. Thus, the new values of SDPE can be written as

$$
\begin{aligned}
\mathrm{SDPE}_{\mathrm{O}}^{+} &= \mathrm{SDPE}_{\mathrm{O}} + \gamma, \\
\mathrm{SDPE}_{\mathrm{E}}^{+} &= \mathrm{SDPE}_{\mathrm{E}} + \gamma,
\end{aligned}
$$

where $\mathrm{SDPE}_{\mathrm{O}}^{+}$ and $\mathrm{SDPE}_{\mathrm{E}}^{+}$ are the values of $\mathrm{SDPE}_{\mathrm{O}}$ and $\mathrm{SDPE}_{\mathrm{E}}$ respectively when the market sales structure is misidentified, and $\mathrm{SDPE}_{\mathrm{O}}$ and $\mathrm{SDPE}_{\mathrm{E}}$ are used herein to represent the values of

SDPE with no misidentification. Hence, ΔSDPE becomes

$$
\begin{aligned}
\Delta\text{SDPE} &= \frac{\text{SDPE}_\text{O}^+ - \text{SDPE}_\text{E}^+}{\text{SDPE}_\text{O}^+} \\
&= \frac{\text{SDPE}_\text{O} - \text{SDPE}_\text{E}}{\text{SDPE}_\text{O} + \gamma}.
\end{aligned}
\tag{19}
$$

Eq. 19 reveals that if misidentification exists (i.e. the value of γ is nonzero), the value of ΔSDPE becomes smaller. This result suggests that the values of ΔSDPE we have presented might be biased downward. Eq. 19 also highlights the importance of proper identification of the market sales structure. If the value of γ becomes large, the benefit of information sharing (i.e. the value of ΔSDPE) will disappear.

C On the values of θ

If SuperStore exploits a different value of ρ, namely $\acute{\rho}$ where $\acute{\rho} \neq \rho$, for its MMSE forecast, the ORDER (Y_t) series will become

$$
\begin{aligned}
Y_t &= X_t + (S_t - S_{t-1}) \\
&= X_t + (\hat{X}_{t+1} - \hat{X}_t) \\
&= X_t + (d + \acute{\rho}(X_t - d) - (d + \acute{\rho}(X_{t-1} - d))) \\
&= d + \rho(X_{t-1} + \acute{\rho}X_{t-1} - \acute{\rho}X_{t-2} - d) + (1 + \acute{\rho})\varepsilon_t - \acute{\rho}\varepsilon_{t-1} \\
&= d + \rho(Y_{t-1} - d) + (1 + \acute{\rho})\varepsilon_t - \acute{\rho}\varepsilon_{t-1} \\
&= d + \rho(Y_{t-1} - d) + \acute{\nu}_t - \acute{\theta}\acute{\nu}_{t-1},
\end{aligned}
\tag{20}
$$

where $\acute{\nu}_t = (1 + \acute{\rho})\varepsilon_t$ and $\acute{\theta} = \acute{\rho}/(1 + \acute{\rho})$. Eq. 20 suggests that even if SuperStore uses $\acute{\rho}$, the ORDER series (Y_t) has still an ARMA(1, 1) structure and retains the true values of d, ρ and ε_t, however, the use of $\acute{\rho}$ affects the values of θ. These phenomena are actually observed in Table 4. Therefore, a potential reason of the differences between θ and $\hat{\theta}_\text{O}$ in Table 4 might be that the forecasting method actually used by SuperStore is not exactly same as the method we have used herein. However, the impact of this difference on our final conclusions might be negligible since the value of $\hat{\theta}_\text{O}$ does not directly affect the accuracy of $\hat{Y}_{t+1}$ (see, Eq. 4 and Eq. 6), and the value differences between θ and $\hat{\theta}_\text{O}$ are small, especially for Low and High volume items (see, Table 4).

References

Alwan, L. C., J. J. Liu and D. Yao (2003). Stochastic characterization of upstream demand process in a supply chain. *IIR Transactions* **35**, 207–219.

Aviv, Y. (2002). Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand. *Manufacturing & Service Operations Management* **4**(1), 55–74.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel (1994). *Time Series Analysis: Forecasting and Control*. 3rd ed. Prentice Hall. Englewood Cliffs, NJ.

Christopher, M. (1998). *Logistics and Supply Chain Management: Strategies for reducing cost and improving service*. 2nd ed. Financial Times Prentice Hall.

Gaur, V., A. Giloni and S. Seshadri (2005). Information sharing in a supply chain under ARMA demand. *Management Science* **51**(6), 961–969.

Gavirneni, S., R. Kapuscinski and S. Tayur (1999). Value of information in capacitated supply chain. *Management Science* **45**(1), 16–24.

Geary, S., S. M. Disney and D. R. Towill (2006). On bullwhip in supply chains - historical review, present practise and expected future impact. *International Journal of Production Economics* **101**, 2–18.

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 nonstationary demand process. *Manufacturing & Service Operations Management* **1**, 50–61.

Hosoda, T. and S. M. Disney (2006). On variance amplification in a three-echelon supply chain with minimum mean square error forecasting. *OMEGA, The International Journal of Management Science* **34**(4), 344–358.

Hurvich, C. M. and C. Tsai (1989). Regression and time series model selection in small samples. *Biometrika* **76**(2), 297–307.

Lee, H. L, K. C. So and C. S. Tang (2000). The value of information sharing in a two-level supply chain. *Management Science* **46**(5), 626–643.

Naim, M. M., P. Childerhouse, S. M. Disney and D. R. Towill (2002). A supply chain diagnostic methodology: Determining the vector of change. *Computers & Industrial Engineering* **43**, 135–157.

Potter, A., C. Lalwani, T. Hosoda and H. Al-Kaabi (2005). Vendor management inventory in a grocery supply chain: What are the benefits? The 10th International Symposium on Logistics, Lisbon, Portugal.

Potter, A., M. M. Naim and S. M. Disney (2004). Utilising a production system archetype to attenuate bullwhip in a grocery supply chain. The 8th International Symposium on Logistics, Seville, Spain.

Raghunathan, S. (2001). Information sharing in a supply chain: A note on its value when demand is nonstationary. *Management Science* **47**(4), 605–610.

Zhang, X. (2004). Evolution of ARMA demand in supply chains. *Manufacturing & Service Operations Management* **6**(2), 195–198.