Hosoda, T. and Disney, S.M., (2012), "On the replenishment strategy when the market demand information is lagged in a supply chain", International Journal of Production Economics, Vol. 135, No. 1, pp458–467. DOI: 10.1016/j.ijpe.2011.08.022.

On the replenishment policy when the market demand information is lagged

Takamichi Hosoda

Graduate School of International Management, Aoyama Gakuin University, Shibuya, Shibuya-ku, Tokyo, 150-8366, Japan

Stephen M. Disney

Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff CF10 3EU, Wales, UK.

Abstract

We consider a situation where the most up-to-date information on the market demand and the inventory levels is not available to a replenishment decision maker in a single echelon of a supply chain. The objective of the decision maker is to minimise the sum of the inventory and the production costs. An intuitively attractive strategy under this setting might be to reduce the information time lag as much as possible by utilising information technologies such as RFID. We call this strategy the Time lag Elimination Strategy (TES). However, this course of action requires investment in information systems and will incur a running cost. We propose an alternative strategy that has similar economic consequences as the TES strategy, but it does not require new information systems. We call this strategy the Controlling Dynamics Strategy (CDS). The benefit coming from CDS is quantified and is compared to that from TES. We also quantify the benefits gained from the combined use of these two strategies. A new ordering policy is introduced that is easy to implement without any forecasting systems and can reduce the production cost significantly.

Key words: Time-lag, information delay, the order-up-to policy, the base stock policy, AR(1) process

⋆ Corresponding author Takamichi Hosoda.

Email addresses: Hosoda@gsim.aoyama.ac.jp (Takamichi Hosoda), DisneySM@cardiff.ac.uk (Stephen M. Disney).

Hosoda, T. and Disney, S.M., (2012), "On the replenishment strategy when the market demand information is lagged in a supply chain", International Journal of Production Economics, Vol. 135, No. 1, pp458–467. DOI: 10.1016/j.ijpe.2011.08.022.

1 Introduction

The availability of the real-time information in a supply chain has generally been promoted as a necessity for good performance. For instance, Johnson and Mena (2008) argue that the use of the real-time data is a key factor for supply chain success. Therefore, obtaining real-time information and sharing it among players (e.g. retailers and manufacturers) in a supply chain can sometimes be the major motivation for implementing expensive supply chain planning systems (Kobayashi *et al.*, 2003) and ERP systems (Kelle and Akbulut, 2005).

These articles, at the same time, imply that there are still many actual instances where real-time information is not available for supply chains to use. In other words, it might be quite usual that time lags exist in the information flow [1] . For example, van der Vorst *et al.* (1998) and Wang *et al.* (2007) respectively describe cases with time lags in the information flow in grocery and construction supply chains. Both advocate that eliminating time lags can bring significant benefits to a supply chain.

Recently, the topic of the lagged information problem in a supply chain is highlighted in more theoretical papers. For example, Zhang (2005) and Miyaoka and Hausman (2004) assume that the order-up-to level is updated every time period by using old forecast information and quantify the impact of such non-optimum forecasting on a supply chain cost. In their model, however, the latest market demand information is exploited when the order quantity is determined. Bensoussan *et al.* (2006) and Bensoussan *et al.* (2007) investigate the optimum ordering policy when inventory information is lagged. They conclude that an Order-Up-To (OUT) type policy is an optimum ordering policy for this problem when only inventory costs are present.

A simple but fundamental question emerges: "Is obtaining real-time information the only strategy to deal with a time lag in the information flow in a supply chain?" We agree that to reduce the time lags as much as possible (by introducing information technologies, for example) is a worthy strategy to be considered by a supply chain redesign. We call this strategy the *Time-lag Elimination Strategy* (TES). However, this course of action may require investment in information technology and may also incur running and maintenance costs.

Here, we also propose an alternative strategy that does not require a large amount of investment but can improve the supply chain performance as much as the TES can. In other words, this new strategy enables the supply chain to reduce costs without removing time lags in the information flow. This strategy

[1] In this paper, "lag" is used almost interchangeably with "delay". Each term represents the shift of the information in time.

is called the *Controlling Dynamics Strategy* (CDS). In this paper, the benefits coming from the CDS is quantified and is compared to that from the TES.

The benefit of manipulating the dynamics of a supply chain has been studied by many researchers (see, for example, Disney and Towill, 2003; Hosoda and Disney, 2006; Gaalman and Disney, 2006). In Disney and Towill (2003), the APIOBPCS model (John *et al.*, 1994) is exploited (APIOBPCS is an acronym for Automatic Pipeline feedback Inventory and Order-Based Production Control System). A simple description of the APIOBPCS is the OUT policy with the exponential smoothing forecasting and proportional feedback controllers. In the APIOBPCS model, three different parameters are used to control the supply chain dynamics, α, T_i and T_w. α is a smoothing parameter used in the exponential smoothing forecasts, T_i is a proportional controller in the net stock feedback loop and T_w is a proportional controller in the work-in-progress feedback loop. Using a single level supply chain model, Disney and Towill (2003) show that the Bullwhip Effect can be reduced. Hosoda and Disney (2006) use a replenishment rule that is slightly different from Disney and Towill (2003). The model considers a two-level supply chain and conditional expectation of the market demand is used for the forecasts, instead of exponential smoothing. In addition, only a single proportional controller is used in the OUT policy. The standard deviation of the net stock levels at each level of the supply chain is used as an indicator of the inventory cost. It is concluded that costs can be reduced by 10% by manipulating the dynamics of the replenishment rule.

Gaalman and Disney (2006) consider a single level supply chain case and use state space techniques to study the ARMA(1, 1) demand case. They analyse the trade-off between inventory and order variances. The research presents another method of controlling the dynamics of a supply chain. A gain is applied to the conditional expectation forecasting method together with a feedback controller within the ordering policy. The modified value of the forecast increases the flexibility for controlling the dynamics of the supply chain and allows for improved performance.

Using a serially linked two level supply chain with an ARMA(1, 1) market demand process, Hosoda and Disney (2009) consider the situation where the supply chain wrongly identifies the demand process as an AR(1) process. They show that demand process mis-specification is not always bad for the supply chain as a whole. Through an intensive numerical analysis, they propose a new ordering policy with a controllable parameter in the forecast (β) and quantify the benefit. They conclude that setting $\beta = 0$ is almost always a good (but not the best) choice in terms of economic performance.

This paper is organised as follows. In the next section the model will be developed and analysed. Then via numerical analysis, properties of the model will be illustrated in Section 3. We conclude in Section 4.

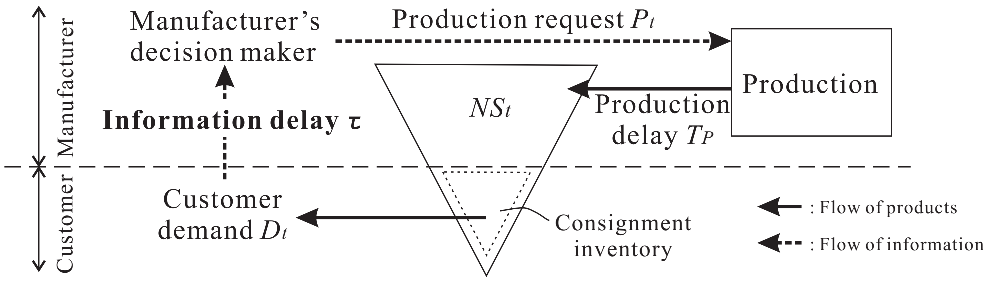


Fig. 1. Schematic of the model

2 The model

Let us use a consignment inventory example to illustrate our model (Fig. 1). Our supply chain model consists of a single player called the manufacturer. In the consignment inventory case, the manufacturer's customer does not place any orders on the manufacturer. Instead, when the need arises, products are picked by the customer from a consignment inventory which is located within the customer's premise.

A sales representative delivers the products from the manufacturer's warehouse to the customer's stock point. It is assumed that the sales representative has easy access to both the customer's inventory and the manufacturer's inventory. Thus, unfilled customer's demand can be satisfied from the manufacturer's on-hand inventory. If the manufacturer's on-hand inventory is not large enough and still some demands are unmet, those demands are backlogged. We can consider the manufacturer's inventory and the consignment inventory as a single pile of inventory.

At the end of each day, by observing the left-over inventory at the customer's stock point, the demand can be observed by the manufacturer's sales representative. This demand information is relayed to the manufacturer's production planner the next day. Thus the demand information that the manufacturer's production planner uses is delayed by one period. There also might be some additional time lags due to the information transaction within the organisations.

The sequence of events when the time lag $\tau = 1$ is assumed to be as follows (Fig. 2). At period $t-1$, the demand (D_{t-1}) occurs and is filled immediately with the consignment inventory at customer's stock point. This demand (D_{t-1}) is observed at the end of the period by a manufacturer's sales representative who visits the customer's stock point to refill the inventory. At the beginning of period t, after a constant production delay (T_P), the previous production request P_{t-T_P-1} is completed, assuming that the infinite raw material is available for the production. Then, at the end of t, the manufacturer's decision maker places a production request, P_t, based on the knowledge of the lagged

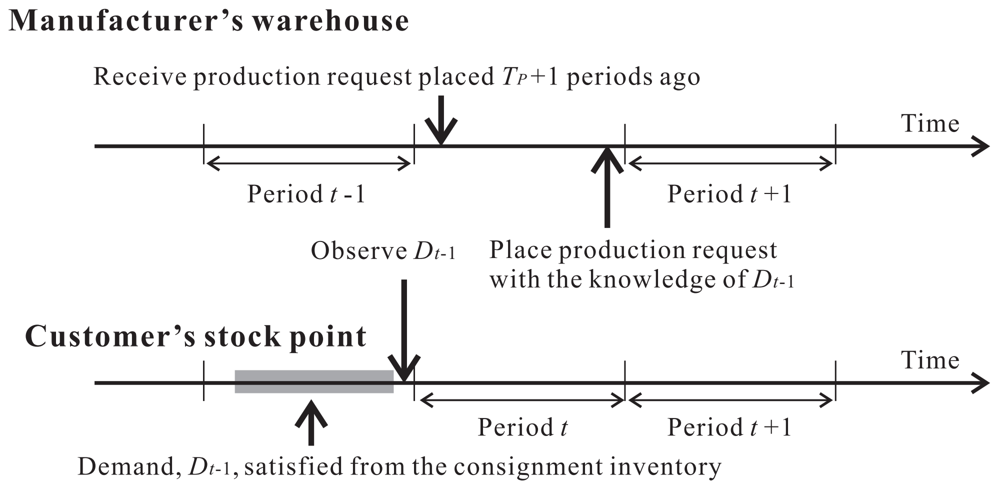


Fig. 2. Sequence of events when $\tau = 1$

demand information, D_{t-1}. Note that $\{\tau, T_P\} \in \aleph_0$.

The manufacturer incurs both an inventory cost and a production cost. Based on the end period net stock level (NS_t), an inventory cost will be charged every time period. A holding cost per unit of on-hand inventory (h) and a backlog cost per unit of unmet demand (b) are used. The production cost is determined by the volume of the production request (P_t) and the standard capacity of the production line, G. Every time period, the manufacturer incurs the fixed cost when P_t is equal to or less the production capacity G. If P_t is greater than G, the manufacturer incurs the fixed cost plus an over-time cost per period for each item produced over capacity G. In this research, it is assumed that the manufacturer's objective is to minimise the sum of the inventory cost and the production cost, subject to the time lags in the information flow.

Before moving on to the next section, let us consider the inventory balance equation. Under the assumption that no time lags in the information flow exists, the manufacturer should have the following inventory balance equation at time period t, $NS_t = NS_{t-1} + P_{t-T_P-1} - D_t$, where NS_t is the net stock level of the manufacturer at t. However, we assume that the manufacturer does not know D_t due to the time lag, but only $D_{t-\tau}$. On the other hand, P_{t-T_P-1} is known to the manufacturer, since it is generated by the manufacturer itself. Note that if in addition to P_{t-T_P-1}, the manufacturer also knows NS_t and NS_{t-1}, then the manufacturer could determine D_t using the inventory balance equation, even though the manufacturer does not observe D_t directly. Thus, we should also assume that the manufacturer does not know $\{NS_t, NS_{t-1}, \ldots, NS_{t-\tau+1}\}$ either when $\tau > 0$.

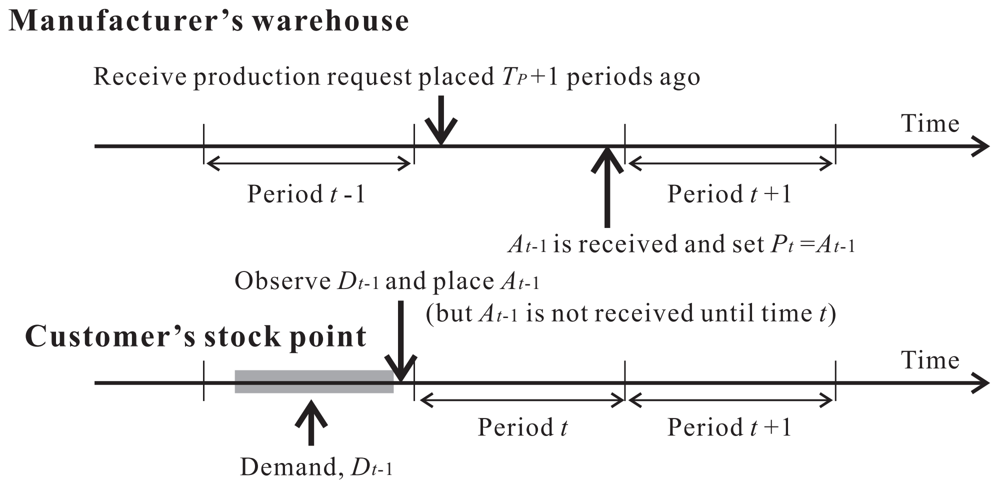


Fig. 3. Relationship between $A_{t-\tau}$ and P_t when $\tau = 1$

2.1 The market demand model

It is assumed that the market demand observed by the manufacturer follows a mean-centred first-order autoregressive process (AR(1)),

$$D_{t-\tau} = \mu + \rho(D_{t-\tau-1} - \mu) + \varepsilon_{t-\tau}, \tag{1}$$

where $D_{t-\tau}$ is the non-negative market demand at time $t - \tau$, μ is the mean of the demand process, ρ is the autoregressive parameter, restricted to $|\rho| < 1$ to ensure the stability of the process. $\varepsilon_{t-\tau}$ is an i.i.d. normally distributed random variable with a mean of zero and a variance of σ_ε^2, $N(0, \sigma_\varepsilon^2)$ at time $t-\tau$. The variance of an AR(1) process is given by $Var[D] = \sigma_\varepsilon^2/(1-\rho^2)$, where $Var[\cdot]$ denotes the variance. Detailed characteristics of an AR(1) process are discussed in Box *et al.* (2008). An AR(1) demand assumption is quite usual in research and actually some real demands can be represented by this model (see Lee *et al.* (2000) and Hosoda *et al.* (2008), for example).

2.2 The order-up-to policy with time lag in the information flow

In this section, we will introduce the OUT policy for the case where the demand information is lagged. We refer to Johnson and Thompson (1975) for a particularly insightful explanation on the OUT policy without time lags in information flow. The situation considered in this research can be restated as follows (see Fig. 3): At the end of period $t-\tau$, a tentative production request, $A_{t-\tau}$, is placed with the knowledge of $D_{t-\tau}$ (by a production planning system, for example). However, due to the time-lag, $A_{t-\tau}$ will be acknowledged (by a production site manager, for example) τ periods later. At the end of t, $A_{t-\tau}$

is recognised and considered as the production request, P_t. Hence, we have $P_t = A_{t-\tau}$.

Let us use $IP^+_{t-\tau}$, the inventory position at $t-\tau$ right after $A_{t-\tau}$ is determined. The inventory position at $t-\tau$ is the sum of the net stock level at $t-\tau$, $NS_{t-\tau}$, and the total of on-orders, $\{A_{t-\tau}, A_{t-\tau-1}, \ldots, A_{t-\tau-T_P}\}$. All these on-orders will become the manufacturer's on-hand inventory during the period $(t-\tau, t+T_P+1]$. However, whatever ordering policy is used, we will always have the following relationship between NS_{t+T_P+1} and $IP^+_{t-\tau}$:

$$NS_{t+T_P+1} = IP^+_{t-\tau} - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}. \tag{2}$$

Note that the above equation reflects the impact of the work-in-progress, $\{A_{t-\tau}, A_{t-\tau-1}, \ldots, A_{t-\tau-T_P}\}$, since $IP^+_{t-\tau}$ includes it.

Generally, for a random variable x, we may obtain the variance of x by using the following expression: $Var[x] = E[(x-\mu)^2]$, where $E[\cdot]$ is the expected value and $\mu = E[x]$. Let $E[IP^+_{t-\tau} - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}]$ denote the *Target Net Stock level* (*TNS*, hereafter) at the end of $t+T_P+1$. From Eq. 2, it is easy to see that *TNS* is also equal[2] to $E[NS]$. With the aid of *TNS*, we can obtain the variance of the net stock level from Eq. 2.

$$\sigma^2_{NS} = E\left[\left(IP^+_{t-\tau} - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i} - TNS\right)^2\right], \tag{3}$$

where σ^2_{NS} is the variance of the net stock levels. Let $\hat{D}^*_{t-\tau}$ be the conditional expectation of the demand over $\tau+T_P+1$ periods given $D_{t-\tau}$. Thus, $\hat{D}^*_{t-\tau}$ can be written as $\hat{D}^*_{t-\tau} = E_{D\tau}\left[\sum_{i=1}^{\tau+T_{P_s}+1} D_{t-\tau+i}\right]$ and $E_{D\tau}$ is the conditional expectation $E[\cdot|D_{t-\tau}]$. Using knowledge of $D_{t-\tau}$ (Eq. 1), a closed form expression for $\hat{D}^*_{t-\tau}$ can be obtained,

$$\hat{D}^*_{t-\tau} = (\tau + T_P + 1)\mu + \rho\,\frac{1-\rho^{\tau+T_P+1}}{1-\rho}(D_{t-\tau} - \mu).$$

Using $\hat{D}^*_{t-\tau}$, we can modify Eq. 3 to yield

[2] The subscript in $E[NS_{t+T_P+1}]$, $t+T_P+1$, is ignored since $E[NS] = TNS$ is time-independent.

$$\sigma_{NS}^2 = E_{D\tau}\left[\left(\left(IP_{t-\tau}^+ - \hat{D}_{t-\tau}^* - TNS\right) + \left(\hat{D}_{t-\tau}^* - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}\right)\right)^2\right]. \quad (4)$$

Furthermore, since $E_{D\tau}\left[\hat{D}_{t-\tau}^* - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}\right] = 0$, Eq. 4 can be simplified to

$$\sigma_{NS}^2 = E_{D\tau}\left[\left(IP_{t-\tau}^+ - \hat{D}_{t-\tau}^* - TNS\right)^2\right] + E_{D\tau}\left[\left(\hat{D}_{t-\tau}^* - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}\right)^2\right]. \quad (5)$$

The first term in RHS of Eq. 5 can be eliminated by setting $IP_{t-\tau}^+ = \hat{D}_{t-\tau}^* + TNS$. On the other hand, the second term in RHS will not be zero since the value of $\sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}$ is unknown to the decision maker and in most practical cases, it is difficult to obtain $\hat{D}_{t-\tau}^* = \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}$ with only knowledge of $\{D_{\tau-t}, D_{\tau-t-1}, \dots\}$. The minimum value of σ_{NS}^2, σ_{NS}^{*2}, is achieved when $IP_{t-\tau}^+ = \hat{D}_{t-\tau}^* + TNS$ and is given by

$$\begin{aligned}\sigma_{NS}^{*2} &= E_{D\tau}\left[\left(\hat{D}_{t-\tau}^* - \sum_{i=1}^{\tau+T_P+1} D_{t-\tau+i}\right)^2\right] \\ &= \frac{\left(\begin{array}{l}(\tau+T_P+1)(1-\rho^2)+ \\ \rho(1-\rho^{\tau+T_P+1})(\rho^{\tau+T_P+2}-\rho-2)\end{array}\right)}{(1-\rho)^2(1-\rho^2)}\sigma_\varepsilon^2.\end{aligned}$$

Property 1 *σ_{NS}^{*2} is increasing in τ and T_P.*

Proof 1 *Let $L = \tau + T_P + 1$ and*

$$\sigma_{NS}^{*2}[L] = \frac{\left(L(1-\rho^2) + \rho(1-\rho^L)(\rho^{L+1}-\rho-2)\right)}{(1-\rho)^2(1-\rho^2)}\sigma_\varepsilon^2.$$

Then as

$$\sigma_{NS}^{*2}[L+1] - \sigma_{NS}^{*2}[L] = \frac{(\rho^{L+1}-1)^2}{(\rho-1)^2}\sigma_\varepsilon^2 > 0,$$

it is easy to see that σ_{NS}^{*2} is increasing in τ and T_P. □

By following the sequence of events and the definition of $IP_{t-\tau}^+$, $A_{t-\tau}$ and P_t can be expressed as

$$\begin{aligned} A_{t-\tau} = P_t &= IP^+_{t-\tau} - IP^-_{t-\tau} \\ &= IP^+_{t-\tau} - (IP^+_{t-\tau-1} - D_{t-\tau}), \end{aligned} \tag{6}$$

where $IP^-_{t-\tau}$ is the inventory position at $t-\tau$ just before $A_{t-\tau}$ is made. Substituting $IP^+_{t-\tau} = \hat{D}^*_{t-\tau} + TNS$ into Eq. 6 yields

$$P_t = F_{\text{OUT}}(D_{t-\tau} - D_{t-\tau-1}) + D_{t-\tau}, \tag{7}$$

where

$$F_{\text{OUT}} = \rho(1-\rho^{\tau+T_P+1})/(1-\rho). \tag{8}$$

Eq. 7 represents the OUT policy when the demand information is lagged τ time periods. The following set of formulae is identical to Eq. 7 and might be easier for the most readers to recognise.

$$\begin{cases} P_t = D_{t-\tau} + (S_t - S_{t-1}) \\ S_t = \hat{D}^*_{t-\tau} + TNS, \end{cases} \tag{9}$$

where S_t represents the OUT level at t. Furthermore, by using Eq. 1, Eq. 7 can be rewritten as

$$P_t = \mu + (\rho + \rho F_{\text{OUT}} - F_{\text{OUT}})(D_{t-\tau-1} - \mu) + (1 + F_{\text{OUT}})\varepsilon_{t-\tau}.$$

Using the characteristic that $D_{t-\tau-1}$ and $\varepsilon_{t-\tau}$ are mutually independent, the variance of P_t, σ^2_P, is

$$\begin{aligned} \sigma^2_P &= E\left[\left((\rho + \rho F_{\text{OUT}} - F_{\text{OUT}})(D_{t-\tau-1} - \mu) + (1 + F_{\text{OUT}})\varepsilon_{t-\tau}\right)^2\right] \\ &= (\rho + \rho F_{\text{OUT}} - F_{\text{OUT}})^2\, Var[D] + (1 + F_{\text{OUT}})^2\, Var[\varepsilon] \\ &= \frac{(\rho + \rho F_{\text{OUT}} - F_{\text{OUT}})^2}{1-\rho^2}\sigma^2_\varepsilon + (1 + F_{\text{OUT}})^2\,\sigma^2_\varepsilon \\ &= \frac{2F_{\text{OUT}}(F_{\text{OUT}}+1)(1-\rho)+1}{1-\rho^2}\sigma^2_\varepsilon. \end{aligned} \tag{10}$$

2.3 A generalised order-up-to policy with time lag in the information flow

From now, let us consider the more general case of $IP^+_{t-\tau} = \hat{D}_{t-\tau} + TNS$, where $\hat{D}_{t-\tau} = (\tau + T_P + 1)\mu + F(D_{t-\tau} - \mu)$ and F is an arbitrary value. Then Eq. 7 and Eq. 9 become

$$P_t = F(D_{t-\tau} - D_{t-\tau-1}) + D_{t-\tau}, \tag{11}$$

and

$$\begin{cases} P_t = D_{t-\tau} + (S_t - S_{t-1}) \\ S_t = \hat{D}_{t-\tau} + TNS, \end{cases} \tag{12}$$

respectively. We call Eq. 11 and 12 the generalised OUT policy herein. Eq. 5 and Eq. 10 can be rewritten as

$$\sigma_{NS}^2 = \frac{\sigma_\varepsilon^2}{1-\rho^2}\left(F - \rho\frac{1-\rho^{\tau+T_P+1}}{1-\rho}\right)^2 + \sigma_{NS}^{*2}, \tag{13}$$

$$\sigma_P^2 = \frac{2F(F+1)(1-\rho)+1}{1-\rho^2}\sigma_\varepsilon^2. \tag{14}$$

Since both σ_{NS}^2 and σ_P^2 are functions of F, the manufacturer can control the dynamics of its inventory and production request by manipulating the value of F. Therefore, we call F the *strategic parameter* for the manufacturer and assume that the manufacturer is free to choose the value of F to achieve its strategic objective. For example, if the manufacturer wants to minimise the variance (or, standard deviation) of the net stock levels (i.e. $\sigma_{NS}^2 = \sigma_{NS}^{*2}$), the manufacturer should set $F = F_{\text{OUT}}$. In this case, the generalised OUT policy becomes the OUT policy. When we set $F = 0$ in Eq. 11, then $P_t = D_{t-\tau}$, which shows that the generalised OUT policy with $F = 0$ is now identical to the Base Stock (BS) policy. The BS policy is essentially an OUT policy with a constant time-invariant OUT level[3] (see, Hopp and Spearman, 2008). Therefore, the generalised OUT policy embraces both the OUT policy and the BS policy.

Property 2 *When the generalised OUT policy (Eq. 11) is exploited, σ_{NS}^2 (Eq. 13) is minimised at $F = F_{\text{OUT}}$.*

Proof 2 *It is obvious from Eq. 13.*

In terms of the impact of τ on σ_P^2, we have the following properties.

Property 3 *When the generalised OUT policy (Eq. 11) is exploited, σ_P^2 (Eq. 14) is independent of τ, if the value of F is determined independently from τ.*

Proof 3 *It is obvious from Eq. 14.*

[3] This situation also can be achieved by setting $\hat{D}_{t-\tau}$ to a time invariant constant, which results in $S_t = S_{t-1}$. From Eq. 12, we can then see that $P_t = D_{t-\tau}$.

Hosoda, T. and Disney, S.M., (2012), "On the replenishment strategy when the market demand information is lagged in a supply chain", International Journal of Production Economics, Vol. 135, No. 1, pp458–467. DOI: 10.1016/j.ijpe.2011.08.022.

Property 4 *When the generalised OUT policy (Eq. 11) is exploited,*

- σ_P^2 *(Eq. 14) is convex in* F,
- *The production request* $P_t = (D_{t-\tau} + D_{t-\tau-1})/2$ *minimises* σ_P^2. *The minimised* σ_P^2, σ_P^{*2}, *is*

$$\sigma_P^{*2} = \frac{1+\rho}{2(1-\rho^2)}\sigma_\varepsilon^2 < \frac{\sigma_\varepsilon^2}{1-\rho^2} = Var[D].$$

Proof 4 *Since*

$$\frac{\partial \sigma_P^2}{\partial F} = \frac{2(1+2F)\sigma_\varepsilon^2}{1+\rho} \quad \text{and} \quad \frac{\partial^2 \sigma_P^2}{\partial F^2} = \frac{4\sigma_\varepsilon^2}{1+\rho} > 0,$$

σ_P^2 *has a single unique minimum value at* $F = -0.5$ *for all values of* ρ, T_P, τ *and* σ_ε. □

Property 4 means that when the production request is equal to the average of the two most recent demands, the variance of the production request is minimised. In addition it suggests that we can eliminate the Bullwhip phenomena, as $\sigma_P^{*2} < Var[D]$.

Hereafter, the generalised OUT policy with $F = -0.5$ is called the HD policy. In addition to the minimisation of σ_P^2, the HD policy has the following characteristic.

Property 5 *When* $D_{t-\tau}$ *follows an AR(1) process,* P_t *generated by the HD policy is an ARMA(1, 1) process.*

Proof 5

$$\begin{aligned} P_t &= (D_{t-\tau} + D_{t-\tau-1})/2 \\ &= (\mu + \rho(D_{t-\tau-1} - \mu) + \varepsilon_{t-\tau} + \mu + \rho(D_{t-\tau-2} - \mu) + \varepsilon_{t-\tau-1})/2 \\ &= \mu + \rho\left((D_{t-\tau-1} + D_{t-\tau-2})/2 - \mu\right) + (\varepsilon_{t-\tau} + \varepsilon_{t-\tau-1})/2 \\ &= \mu + \rho\left(P_{t-1} - \mu\right) + \zeta_t + \zeta_{t-1}, \end{aligned} \tag{15}$$

where ζ_t *is an i.i.d. random variable and* $\zeta_t = 0.5\varepsilon_{t-\tau}$. *Eq. 15 is an ARMA(1, 1) process with a moving average parameter* $\theta = -1$. *Thus,* P_t *is stationary but not invertible* (Box *et al.*, 2008).

The HD policy is a very easy policy to implement. The order rate for each period is the average of the last two observed demand rates. No new information systems are needed. No complex forecasting has to be completed. The HD policy is a special case of the so-called Minimum Variance Policy of Balakrishnan *et al.* (2004).

Hosoda, T. and Disney, S.M., (2012), "On the replenishment strategy when the market demand information is lagged in a supply chain", International Journal of Production Economics, Vol. 135, No. 1, pp458–467. DOI: 10.1016/j.ijpe.2011.08.022.

2.4 The manufacturer's costs

A inventory cost and a production cost are used to measure the manufacturer's performance. The inventory cost per time unit, C_{NS}, can be described as

$$C_{NS} = E\left[h(NS_t)^+ + b(-NS_t)^+\right],$$

where $(\cdot)^+$ is the maximum operator. If NS_t is positive, the manufacturer incurs a holding cost, hNS_t, at the end of t, and if it is negative, the manufacturer incurs the backlog cost, $b(-NS_t)$, at the end of t. If it is assumed that $h < b$, the value of TNS which is equal to $E[NS]$, is a positive but finite value. The optimum value of TNS that minimises the inventory cost per time period can be obtained through the newsvendor problem procedure and is $z_{NS}\sigma_{NS}$, where $z_{NS} = \Phi^{-1}[b/(b+h)]$ for the standard normal distribution Φ. Thus, the minimum inventory cost per time period is $C^*_{NS} = (h+b)\,\phi[z_{NS}]\sigma_{NS}$. Here ϕ is the probability density function of the standard normal distribution.

The production cost, C_P, is determined as follows: It is assumed that the manufacturer has a time-invariant constant production capacity, G. G is determined as $G = \mu + s$, where s is the slack capacity above or below the mean demand and is set to minimise C_P. If the production request P_t is equal to or lower than G, then some employees may idle but they are still paid. In this case, the cost for the manufacturer is $u(\mu + s)$, where u is the standard cost rate to produce one unit of the product. When P_t is greater than G, an over-time shift is used to meet the production request. The cost to produce a unit in the over-time is w and is assumed that $w > u$. The production cost per time unit can be written as

$$C_P = E\left[u(\mu + s) + w(P_t - (\mu + s))^+\right]. \tag{16}$$

The slack capacity that minimizes C_p, s^*, is given by $z_s\sigma_P$ where $z_s = \Phi^{-1}[(w-u)/w]$. Details of how this derived are shown in Appendix 1. When the capacity G is set to $\mu + s^*$, the production cost becomes

$$C^*_P = u\mu + w\phi\left[z_s\right]\sigma_P. \tag{17}$$

Both C^*_{NS} and C^*_P are linear functions of σ_{NS} and σ_P respectively, and σ_{NS} and σ_P are functions of the strategic parameter F.

Together with Property 1, we can conclude that if minimisation of the inventory cost (C^*_{NS}) is the only concern for the manufacturer, the manufacturer should set $F = F_{\text{OUT}}$ (i.e. the OUT policy) and try to eliminate time lags in the demand information as much as possible. On the other hand, Property 4 suggests that if the manufacturer's concern is only about reducing its production cost (C_P^*), the manufacturer will set $F = -0.5$ (i.e. the HD policy) and may not be interested in reducing the time lags in the demand information, since C_P^* is independent of τ in this case. We may also consider the following objective function:

$$J[F] = C_{NS}^* + C_P^*,$$

subject to a given set of values of $\{\mu, \rho, \sigma_\varepsilon, T_P, \tau, h, b, u, w\}$.

Property 6 *The optimum value of F, F^*, which minimises the value of $J[F]$, exists between -0.5 and F_{OUT}. When ρ satisfies $2\rho^{\tau+T_P+1} - \rho \geq 1$, it is $-0.5 \leq F^* \leq F_{\text{OUT}}$.*

Proof 6 *C_{NS}^* is linear in σ_{NS}^2 (Eq. 13). Thus C_{NS}^* is convex in F and has a unique single minimum at $F = F_{\text{OUT}}$ (Property 2). C_P^* is linear in σ_P (Eq. 17). Hence C_P^* is convex in F and has a unique single minimum at $F = -0.5$ (Property 4). Thus F^* lies between -0.5 and F_{OUT}. $-0.5 \leq F^* \leq F_{\text{OUT}}$ occurs only if $-0.5 \leq F_{\text{OUT}}$. By simplifying $-0.5 \leq$ Eq. 8, we obtain $2\rho^{\tau+T_P+1} - \rho \geq 1$.*

In the next section, we will exploit numerical analysis to quantify the impacts of those two strategies, TES and CDS, on the manufacturer's total cost, $J[F]$.

3 Numerical analysis

In this section, unless otherwise stated, the following values are used for cost parameters: $h = 2$, $b = 50$, $u = 25$, and $w = 50$. The production delay is $T_P = 4$. For the demand process, $\mu = 100$ and $\sigma_\varepsilon = 10$ are assumed. First of all, we will show the impact of F on the supply chain costs. Then we consider the cost benefit from the two strategies.

Fig. 4 shows how the manufacturer can control its dynamics and costs by manipulating the strategic parameter F, when ρ is a positive value. It can be observed that the costs are affected by F, an F^* exists and its value is always $-0.5 \leq F^* \leq F_{\text{OUT}}$. F_{OUT} which always minimises C_{NS}^*, does not minimise $J[F]$. This example suggests that the OUT policy is not optimal for the minimisation of a cost function that includes production cost. Instead, the decision maker is encouraged to choose the value of F to lower $J[F]$.

In Fig. 5, for the base case, we have set $F = F_{\text{OUT}}$, and $\tau = 1$. We then assume the time lag is eliminated (i.e. $\tau = 0$) in the TES. Note that from Eq. 8, F_{OUT} is a function of τ. Thus, the value of F_{OUT} when $\tau = 0$ is different from that

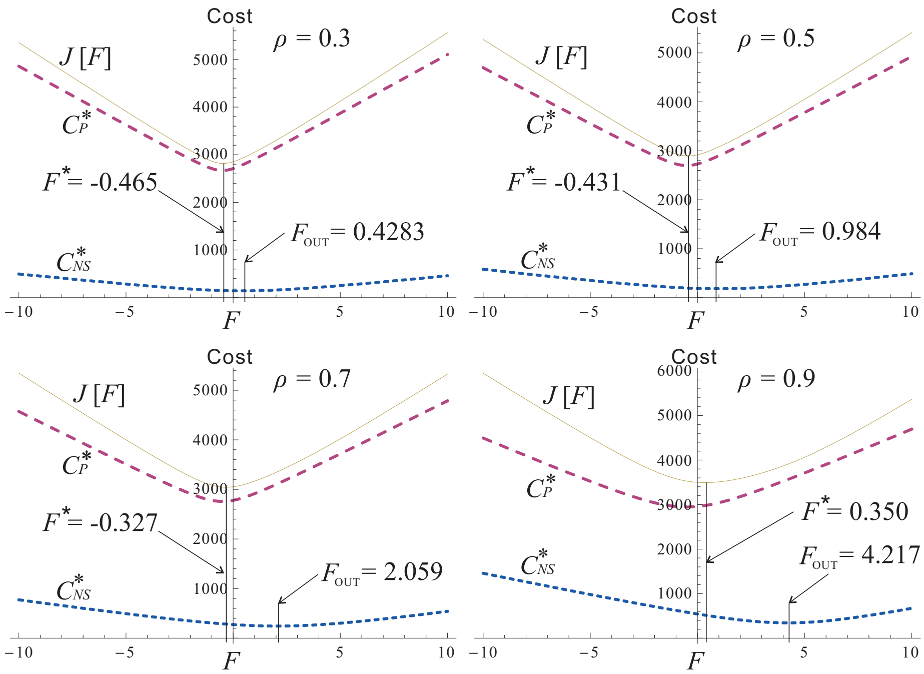


Fig. 4. The impact of F when $\tau = 1$

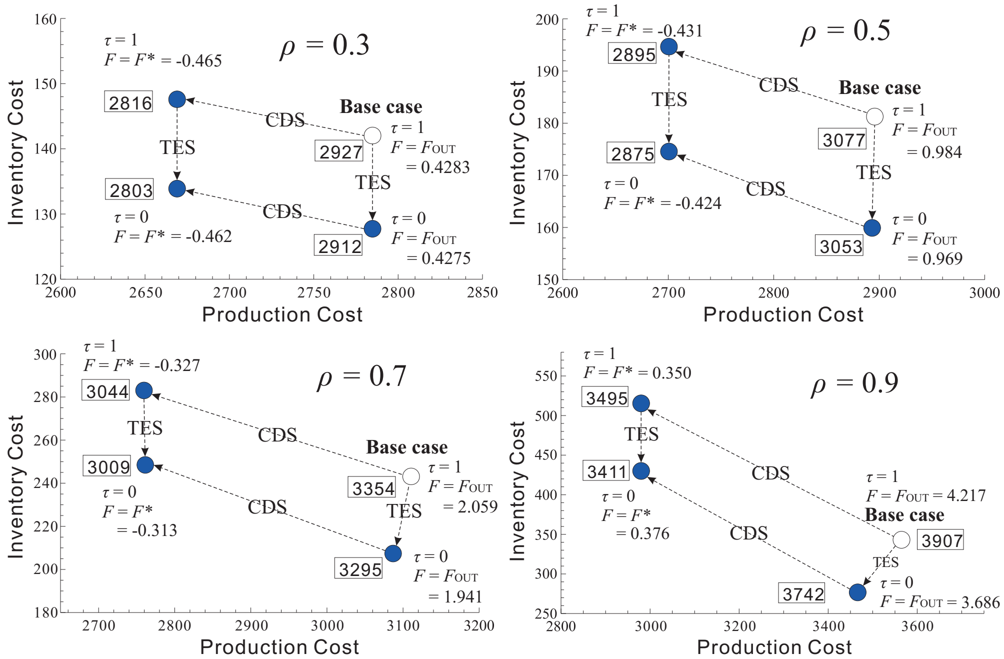
when $\tau = 1$. We also apply the CDS to the base case, there is still a time lag ($\tau = 1$), but we have set $F = F^*$ to minimise $J[F]$. Finally, we also calculate the value of $J[F]$ when the manufacturer exploits those two strategies at the same time, namely $\tau = 0$ and $F = F^*$.

It is observed that TES has a significant impact on the inventory cost most of the time. On the other hand, the CDS never improves inventory cost since F is not equal to F_{OUT} any more when the CDS is applied. However, CDS significantly reduces the production cost and the saving is large enough to compensate for the increase of the inventory cost, since $J[F]$ always becomes lower after the CDS is applied. Furthermore, when TES and CDS are exploited at the same time, $J[F]$ shows the minimum cost.

From a practical point of view, however, it probably quite hard to find F^* as it requires accurate knowledge of the actual demand process (Hosoda and Disney, 2009). In such a case, it could be interesting if the cost performance given by the BS policy or the HD policy is competitive as these policies do not require the precise knowledge and specification of the demand process to determine the order rate. Let us use TC as an indicator of the weighted sum of the production cost and the inventory cost:

$$TC = V\sigma_P + (1 - V)\sigma_{NS},$$

where $V = w\phi[z_s]/(w\phi[z_s] + (h + b)\phi[z_{NS}])$ and $0 < V < 1$. Note that $u\mu$ in C_P^* (Eq. 17) is ignored in TC as the value of $u\mu$ is always constant, as



TES: Time lag Elimination Strategy
CDS: Controlling Dynamics Strategy
$\boxed{J[F]}$

Fig. 5. Performance improvement by each strategy

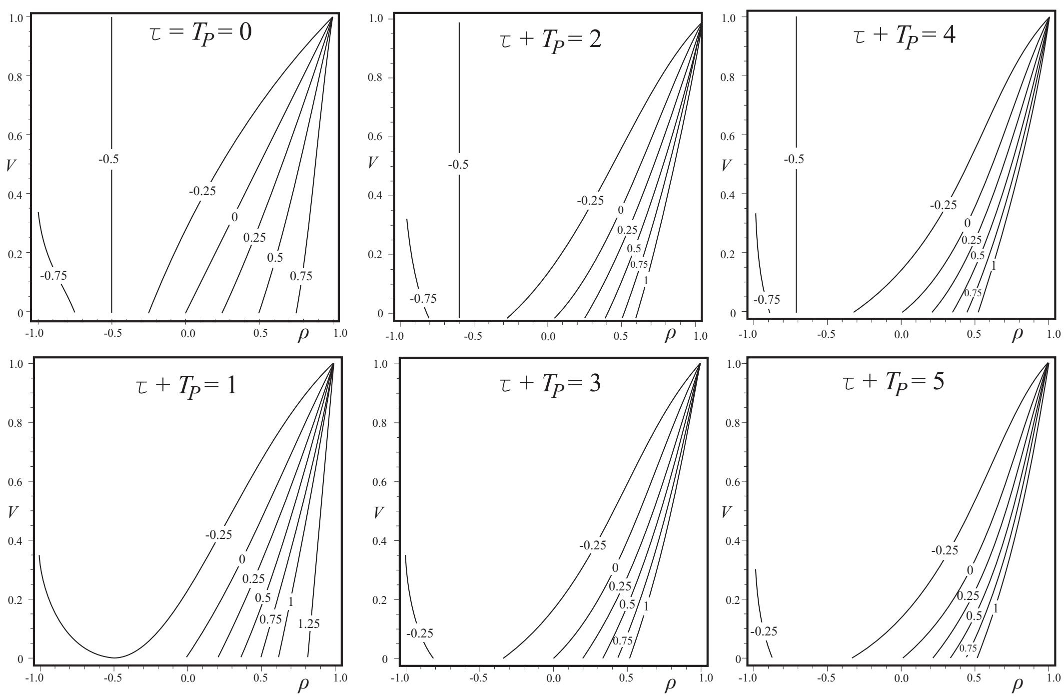
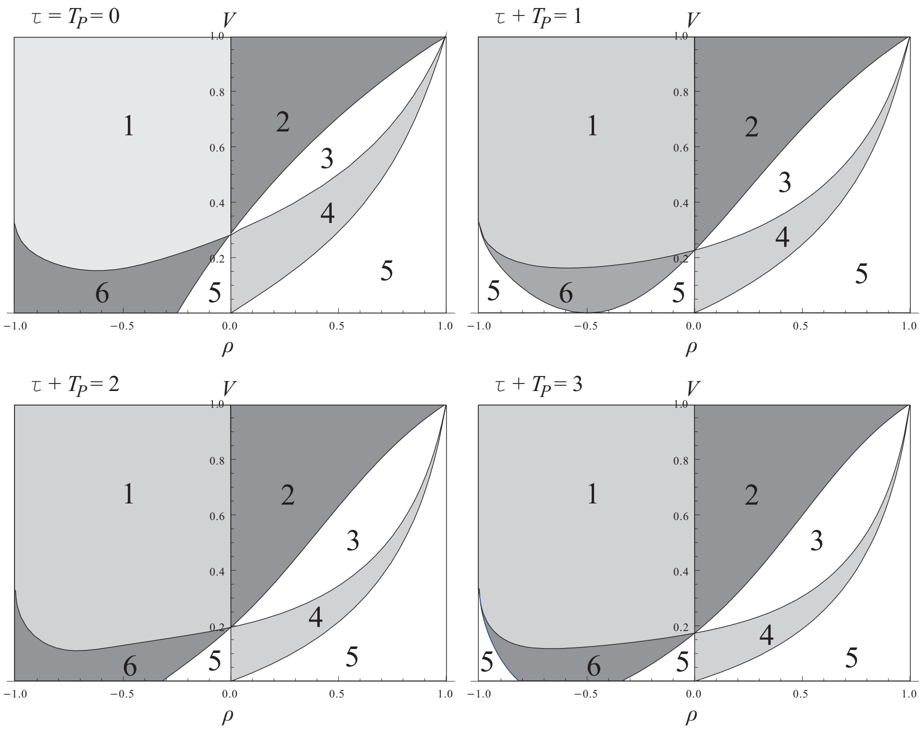


Fig. 6. The optimal F^* in ρ -V plane



Region 1: HD $\leq$ OUT $\leq$ BS, Region 2: HD $\leq$ BS $\leq$ OUT, Region 3: BS $\leq$ HD $\leq$ OUT, Region 4: BS $\leq$ OUT $\leq$ HD, Region 5: OUT $\leq$ BS $\leq$ HD, Region 6: OUT $\leq$ HD $\leq$ BS.

Fig. 7. Cost comparison of ordering policies in ρ -V plane

it is independent of F and disappears when the derivative are taken, when optimising F. By calculating TC as a convex combination of a single weighting factor V, we may obtain insights into how the policy behave for different ratios of h, b, and w. Recall $z_s = \Phi^{-1}[(w-u)/w]$, thus u still influences TC. Note that $TC \times (w\phi[z_s] + (h+b)\phi[z_{NS}]) + u\mu = J[F]$, thus there is no fundamental difference between TC and $J[F]$ - understanding one allows you to understand the other.

Differentiating TC w.r.t. F results in

$$\frac{\partial TC}{\partial F} = \left(\frac{V(1+2F)}{(1+\rho)\sqrt{\frac{2F(1+F)(\rho-1)-1}{\rho^2-1}}} + \frac{(V-1)\left(F(\rho-1)+\rho-\rho^{\tau+T_P+2}\right)}{\sqrt{(1-\rho^2)\left(\tau+T_P+1+F^2(\rho-1)^2-2\rho-\rho^2(\tau+T_P+1-2\rho^{\tau+T_P})-2F(\rho-1)\rho(\rho^{\tau+T_P+1}-1)\right)}} \right),$$

which can be set to zero and solved for the stationary points. There are four solutions, and thus the expression for F^* is rather lengthy. However, we may enumerate it as illustrated in Fig. 6. In Fig. 6 which seems to suggest that $-0.5 \leq F^* \leq F_{\text{OUT}}$ for all ρ, V, τ and T_P when $\tau + T_P$ is odd. However $F_{\text{OUT}} \leq F^* < -0.5$ for ρ near -1 when $\tau + T_P$ is even.

Fig. 7 shows the results when $\tau + T_P$ varies from zero to three, $|\rho| < 1$ and

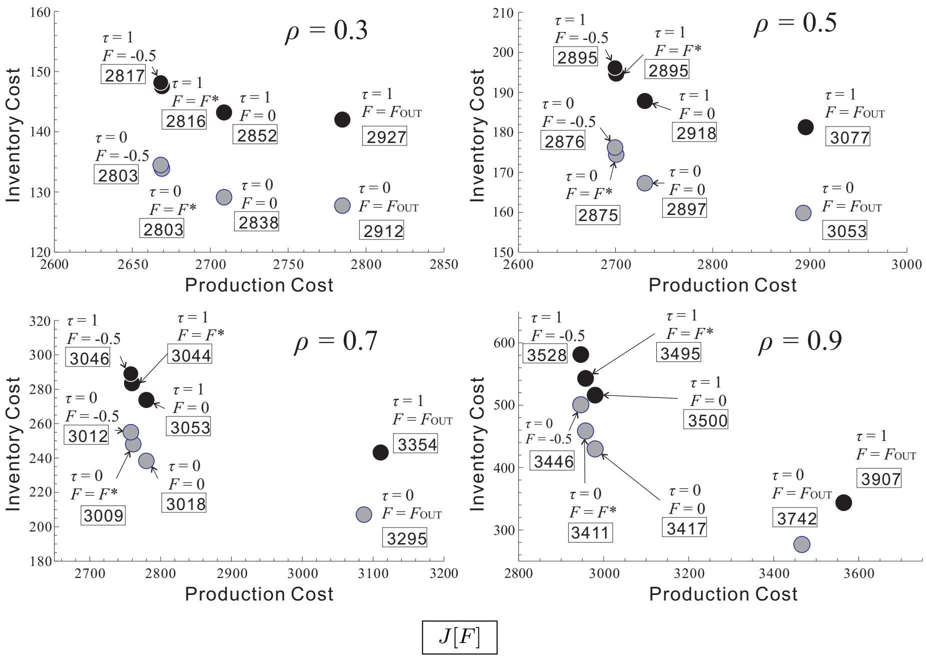


Fig. 8. Cost comparison when $F = F_{\mathrm{OUT}}$, $F = F^{*}$, $F = 0$ and $F = -0.5$

$0 < V < 1$. It is found that there are 6 regions depending on the relative magnitude of each policy's TC. Region 1, for example, is where TC and hence $J[F]$ is minimised by the HD policy. The OUT policy outperforms the BS policy though. It is shown that the ordering policies without any forecasts (i.e. the BS policy and the HD policy) dominate the whole region, irrespective of the value of $\tau + T_P$ (see, the region 1, 2, 3 and 4) unless the production costs are relatively unimportant. Furthermore, it can be concluded that if reducing the production cost is the highest priority (i.e. V is near 1), use of the HD policy is always recommended, as the HD policy always yields the lowest value of TC (and $J[F]$) at any values of $\tau + T_P$ and ρ. We believe this is quite an interesting practical finding. To reduce the production cost, simply take the average of the last two observed consecutive demands and use that as the production order. The HD policy does not require any sophisticated knowledge of the structure of the market demand process, forecast or cost structure of the supply chain.

Numerical results of the total cost are shown in Fig. 8 for $h = 2$, $b = 50$, $u = 25$ and $w = 50$. The costs for the OUT policy (i.e. $F = F_{\mathrm{OUT}}$), the BS policy (i.e. $F = 0$) and the HD policy (i.e. $F = -0.5$) are calculated, together with the case where $F = F^{*}$. The HD policy always yields minimum production cost. Most of the cases, the HD policy or the BS policy shows quite good performance in spite of its simple order decision rule as $J[F = -0.5]$ or $J[F = 0]$ is quite close to $J[F = F^{*}]$.

4 Conclusions

It seems that now we have an answer to the research question: “Is obtaining real-time information the only strategy to overcome the time lag in the demand information?” We can conclude that there is a possible alternative strategy, which we term the CDS. We showed that TES is beneficial, especially for the reduction of the inventory cost, but the CDS can reduce the production cost significantly. This suggests that if the major concern of the supply chain is reducing its inventory cost, TES should be exploited, and if it is its production cost, CDS is recommended. We have not, however, shown that CDS is the best possible strategy for dealing with time lags (linear or otherwise), only a beneficial one.

From a practical point of view, there might be significant differences between these two strategies. To complete the TES, the supply chain may need to make a large investment in information systems (such as RFID), and incur a running cost. On the other hand, CDS only needs to modify the value of F used within the OUT policy, and no new hardware or software is necessary. The running cost for CDS might also be negligible. Therefore, the return on investment may be higher for CDS than for TES.

The benefit of the combined use of two strategies has also been quantified. It is shown that although time lags are eliminated thanks to the TES, F_{OUT} cannot minimise the total cost. Therefore it is recommended that the manufacturer manipulates the dynamics of the supply chain by using the strategic parameter F to lower the total cost, irrespective of the existence of the time lag. Allowing F in the generalised OUT policy to be zero results in the BS policy. When $F = -0.5$, a new ordering policy results. We call this the HD policy.

It is shown that the HD policy is a very attractive alternative when the production cost is the major concern to a supply chain. Its cost performance and its ease of implementation would be attractive for many manufacturers.

Finally, we have noticed in our past experience that labor costs are much more dominant than inventory costs in many manufacturing and low volume warehousing contexts. Thus, V is likely to be very near unity. In this case, the HD policy is a good practical, easy to implement policy.

Acknowledgement

We gratefully acknowledge the support by the Japan Society for the Promotion of Science (Grant No. 22530372), the Daiwa Foundation (Grant No.

7590/8132) and Cardiff Business School Overseas Travel Fellowship Scheme for financial assistance in this research. We are also thankful to Professor Gerard Gaalman (University of Groningen, the Netherlands) and two anonymous referees for useful feedback on preliminary versions of this paper.

References

Balakrishnan, A., Geunes, J. and Pangburn, M.S., 2004. Coordinating supply chains by controlling upstream variability propagation. *Manufacturing & Service Operations Management* **6**, 163–183.

Bensoussan, A., Çakanyıldırım, M. and Sethi, S. P., 2006. Optimality of base-stock and (s, S) policies for inventory problems with information delays. *Journal of Optimization Theory and Applications* **130**, 153–172.

Bensoussan, A., Çakanyıldırım, M. and Sethi, S. P., 2007. Optimal ordering policies for inventory problems with dynamic information delays. *Production and Operations Management* **16**, 241–256.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 2008. *Time Series Analysis: Forecasting and Control*. 4th edition. John Wiley & Sons, NJ.

Disney, S.M. and Towill, D.R., 2003. On the bullwhip and inventory variance produced by an ordering policy. *OMEGA, The International Journal of Management Science* **31**, 157–167.

Gaalman, G. and Disney, S.M., 2006. State space investigation of the bullwhip problem with ARMA(1, 1) demand processes. *International Journal of Production Economics* **104**, 327–339.

Hopp, W.J. and Spearman, M.L., 2008. *Factory Physics*. International edition. McGraw-Hill, NJ.

Hosoda, T. and Disney, S.M., 2006. The governing dynamics of supply chains: The impact of altruistic behaviour. *Automatica* **42**, 1301–1309.

Hosoda, T., Naim, M.M., Disney, S.M. and Potter, A., 2008. Is there a benefit to sharing market sales information? Linking theory and practice. *Computers & Industrial Engineering* **54**, 315–326.

Hosoda, T. and Disney, S.M., 2009. Impact of market demand mis-specification on a two-level supply chain. *International Journal of Production Economics* **121**, 739–751.

John, S., Naim, M.M., and Towill, D.R., 1994. Dynamic analysis of a WIP compensated decision support system. *International Journal of Manufacturing System Design* **1**, 283–297.

Johnson, M. and Mena, C., 2008. Supply chain management for servitised products: A multi-industry case study. *International Journal of Production Economics* **114**, 27–39.

Johnson, G.D. and Thompson, H.E., 1975. Optimality of myopic inventory policies for certain dependent demand processes. *Management Science* **21**, 1303–1307.

Kelle, P. and Akbulut, A., 2005. The role of ERP tools in supply chain information sharing, cooperation, and cost optimization. *International Journal of Production Economics* **93-94**, 41–52.

Kobayashi, T., Tamaki, M. and Komoda, N., 2003. Business process integration as a solution to the implementation of supply chain management systems. *Information & Management* **40**, 769–780.

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

Miyaoka, J. and Hausman, W., 2004. How a base stock policy using "stale" forecasts provides supply chain benefits. *Manufacturing & Service Operations Management* **6**, 149–162.

van der Vorst, J.G.A.J., Beulens, A.J.M., de Wit, W. and van Beek, P., 1998. Supply chain management in food chains: Improving performance by reducing uncertainty. *International Transactions in Operational Research* **5**(6), 487–499.

Wang, L., Lin, Y., and Lin, P.H., 2007. Dynamic mobile RFID-based supply chain control and management system in construction. *Advanced Engineering Informatics* **21**, 377–390.

Zhang, X., 2005. Delayed demand information and dampened bullwhip effect. *Operations Research Letters* **33**, 289–294.

Appendix 1

The expected number of products produced in over-time during the unit time period can be written as

$$\frac{1}{\sigma_P\sqrt{2\pi}}\int_0^{+\infty} e^{-\frac{(s+x)^2}{2\sigma_P^2}}x dx = \frac{\sigma_P}{\sqrt{2\pi}}e^{-\frac{s^2}{2\sigma_P^2}} - \frac{s}{2}erfc\left[\frac{s}{\sigma_P\sqrt{2}}\right],$$

where $erfc[\cdot]$ is the complementary error function. The expected production cost per time period, C_P shown in Eq. 16, is rewritten as

$$C_P = u(\mu+s) + w\left(\frac{\sigma_P}{\sqrt{2\pi}}e^{-\frac{s^2}{2\sigma_P^2}} - \frac{s}{2}erfc\left[\frac{s}{\sigma_P\sqrt{2}}\right]\right).$$

Since the second-order differential of C_P w.r.t. s is always positive,

$$\frac{\partial^2 C_P}{\partial s^2} = \frac{we^{-\frac{s^2}{2\sigma_P^2}}}{\sigma_P\sqrt{2\pi}} > 0,$$

we can have the optimal value of the slack capacity, s^*, by setting the first-order differential is equal to zero.

$$\frac{\partial C_P}{\partial s} = u - \frac{w}{2}\left(1 - erf\left[\frac{s}{\sigma_P\sqrt{2}}\right]\right) = 0.$$

Solving the above equation yields the optimal value of the slack capacity, s^*,

$$s^* = z_s\sigma_P,$$

where

$$z_s = \sqrt{2}erf^{-1}\left[\frac{w-2u}{w}\right] = \Phi^{-1}\left[\frac{w-u}{w}\right].$$