Vlerick Leuven Gent Management School

the Autonomous Management School of
Ghent University and Katholieke Universiteit Leuven

Vlerick Leuven Gent Working Paper Series 2007/04

EXPLORING THE BULLWHIP EFFECT

BY MEANS OF SPREADSHEET SIMULATION

ROBERT N. BOUTE

Robert.Boute@vlerick.be

MARC R. LAMBRECHT

EXPLORING THE BULLWHIP EFFECT

BY MEANS OF SPREADSHEET SIMULATION

ROBERT N. BOUTE

Vlerick Leuven Gent Management School

MARC R. LAMBRECHT

KU Leuven

Contact:

Robert Boute

Vlerick Leuven Gent Management School

Tel: +32 16 24 88 96

Fax: +32 16 24 88 00

Email: Robert.Boute@vlerick.be

ABSTRACT

An important supply chain research problem is the bullwhip effect: demand fluctuations increase as one moves up the supply chain from retailer to manufacturer. It has been recognized that demand forecasting and ordering policies are two of the key causes of the bullwhip effect. In this paper we present a spreadsheet application, which explores a series of replenishment policies and forecasting techniques under different demand patterns. It illustrates how tuning the parameters of the replenishment policy induces or reduces the bullwhip effect. Moreover, we demonstrate how bullwhip reduction (order variability dampening) may have an adverse impact on inventory holdings. Indeed, order smoothing may increase inventory fluctuations resulting in poorer customer service. As such, the spreadsheets can be used as an educational tool to gain a clear insight into the use or abuse of inventory control policies and improper forecasting in relation to the bullwhip effect and customer service.

Keywords: Bullwhip effect, forecasting techniques, replenishment rules, inventory fluctuations, spreadsheet simulation

1. INTRODUCTION: TEACHING THE BULLWHIP PROBLEM

The bullwhip effect is a well-known phenomenon in supply chain management. In a single-item two-echelon supply chain, it means that the variability of the orders received by the manufacturer is greater than the demand variability observed by the retailer. This phenomenon was first popularised by Jay Forrester (1958), who did not coin the term bullwhip, but used industrial dynamic approaches to demonstrate the amplification in demand variance. At that time, Forrester referred to this phenomenon as "Demand Amplification". Forrester's work has inspired many researchers to quantify the bullwhip effect, to identify possible causes and consequences, and to suggest various countermeasures to *tame* or reduce the bullwhip effect.

A number of researchers designed games to illustrate the bullwhip effect. The most famous game is the "Beer Distribution Game". This game has a rich history: growing out of the industrial dynamics work of Forrester and others at MIT, it is later on developed by Sterman in 1989. The Beer Game is by far the most popular simulation and the most widely used game in many business schools, supply chain electives and executive seminars. Simchi-Levi et al. (1998) developed a computerized version of the beer game, and several versions of the beer game are nowadays available, ranging from manual to computerized and even web-based versions (e.g. Machuca and Barajas 1997, Chen and Samroengraja 2000, Jacobs 2000).

Beyond the games, real cases are used as teaching tools to introduce and to address the bullwhip effect (Lee et al 2004). The case study Barilla SpA (Hammond 1994), a major pasta producer in Italy, provides vivid illustrations of issues concerning the bullwhip effect. For a long time, Barilla offered special price discounts to customers who ordered full truckload quantities. Such marketing deals created customer order patterns that were highly spiky and erratic. The supply chain costs were so high that they outstripped the benefits from full truckload transportation. The Barilla case was one of the first published cases that supported empirically the bullwhip phenomenon.

Campbell Soup's chicken noodle soup experience (Cachon and Fisher 1997) is another example. Campbell Soup sells products whose customer demand is fairly stable; the consumption doesn't swing wildly from week to week. Yet the manufacturer faced extremely variable demand on the factory level.

4

After some investigation, they found that the wide swings in demand were caused by the ordering practices of retailers. The swing was induced by forward buying. More recent teaching cases that address the bullwhip effect include Kuper and Branvold (2000), Hoyt (2001) and Peleg (2003).

The objective of this paper is to present a spreadsheet application that can be used for educational purposes to illustrate the impact of the replenishment policy and the forecasting technique on the bullwhip effect. It has been recognized that demand forecasting and the type of ordering policy used are among two of the key causes of the bullwhip effect (Lee et al. 1997a). Lee et al. (1997b) provide a mathematical proof that variance amplification takes place when the retailer adjusts his ordering decision based on demand signals. Dejonckheere et al. (2003) demonstrate that the use of "non-optimal" forecasting schemes, such as the exponential smoothing and moving average forecast, always lead to bullwhip, independent of the observed demand pattern. As such, there has been an increasing number of studies devoted to the adverse effects of demand signaling, improper forecasting and the replenishment rule used (e.g. Watson and Zheng 2002).

In this paper we explore a series of replenishment rules (standard and generalized order-up-to policies) and forecasting methods confronted with different demand processes (identically and independently distributed demand and autoregressive demand processes). What often appears to be a rational policy of the decision maker creates tremendous order amplification. We compare our simulation results with the analytical results available in the literature. The spreadsheets are designed in Microsoft Excel so they are user-friendly and easy to understand.

The remainder of this paper is organized as follows. In the next section we present our spreadsheet model. Section 3 analyses the impact of the standard order-up-to policy with different forecasting techniques on the bullwhip effect. Section 4 describes a generalized order-up-to policy which is able to dampen the order variability for any demand process, and we discuss its impact on customer service. Finally we summarize our findings in section 5.

2. DESCRIPTION OF SPREADSHEET MODEL

Our model follows the standard setup of the Beer Distribution Game (Sterman 1989). Each period, we have the following sequence of events: (1) incoming shipments from the upstream decision-maker are received and placed in inventory, (2) incoming orders (demand) are received from the downstream decision-maker and either filled (if inventory is available) or backlogged, and (3) a new order is placed and passed to the upstream echelon. The inventory position is reviewed every Rp periods. The physical lead time equals Tp periods. The total lead time (risk period) is then equal to $L = Rp + Tp$ periods. We analyze inventory and order fluctuations for a single echelon. Extending the analysis to multiple echelons poses no problems.

There are two basic types of inventory replenishment rules: continuous time, fixed order quantity systems on the one hand and periodic review systems on the other. Fixed order quantity systems result in the same quantity (or multiples thereof) of product being ordered at varying time intervals. In periodic systems, a variable amount of product is ordered at regular, repeating intervals. Given the common practice in retailing to replenish inventories frequently (e.g., daily) and the tendency of manufacturers to produce to demand, our spreadsheet application is based on a periodic review policy. Such a policy is optimal when there is no fixed ordering cost and both holding and shortage costs are proportional to the volume of on-hand inventory or shortage (Nahmias 1997, Zipkin 2000).

In a standard periodic review order-up-to policy, the inventory position IP_t is tracked at the end of every review period Rp and compared with an order-up-to (OUT) level S_t. IP_t is the sum of the net stock NS_t and the inventory on order WIP_t. A positive net stock represents inventory on hand (items immediately available to meet demand), whereas a negative net stock refers to a backlog (demand that could not be fulfilled and still has to be delivered). The inventory on order is the work-in-process, or the items ordered but not yet arrived due to the physical lead time. A replenishment order is then placed to raise the inventory position to the order-up-to or base-stock level:

$$O_t = S_t - IP_t . \tag{1}$$

Analogous to the beer game setup, we assume the review period is one period (Rp = 1), which implies that we place an order every period. The order-up-to level covers the (forecasted) average demand during the risk period and a safety stock to buffer higher than expected demands during the same risk period. We define *lead time demand* as the demand during the risk period L, or $D_t^L = \sum_{j=1}^{L} D_{t+j}$.

In the next section we elaborate on this replenishment policy and define several techniques to forecast (lead time) demand. In the remainder of this section we focus on the structure of the spreadsheets. We define three parts: (1) the input section, where the user selects the parameters of the demand process, the replenishment policy and the forecasting method, (2) the simulation over time, where the user can track the calculations how orders are generated, and (3) the output section, where the key performance measures of the simulation are summarized, together with some illustrating graphs. The spreadsheets can be downloaded from

http://www.econ.kuleuven.be/public/NDBAA78/BullwhipExplorer.xls

2.1 Input section

In the input section, the user defines the parameters of the customer demand process and the forecasting technique. The cells of the parameters that can be changed are shaded. We blocked the cells with automatic calculations in the spreadsheets in order to avoid mistakes and miscalculations. The protection can easily be removed using the Unprotect Sheet command (Tools menu, Protection submenu).

We distinguish between an independent and identically distributed (IID) demand process and a first order autoregressive AR(1) demand (Box and Jenkins 1976). We define the demand process as follows:

$$D_t = \overline{D} + \rho\left(D_{t-1} - \overline{D}\right) + \varepsilon_t, \tag{2}$$

where D_t represents the demand in period t, $\overline{D}$ is the average demand, ρ the autocorrelation coefficient and ε_t a normally distributed IID random error with mean 0 and variance $\sigma_\varepsilon{}^2$. The demand variance equals $\sigma_D^2 = \sigma_\varepsilon^2/(1-\rho^2)$. When demand is IID, the autocorrelation coefficient $\rho = 0$. For $-1 < \rho < 0$, the process is negatively correlated and exhibits period-to-period oscillatory behavior.

7

For $0 < \rho < 1$, the demand process is positively correlated which is reflected by a meandering sequence of observations.

The user can select a transportation lag, or physical lead time Tp. This in turn determines the risk period L = Tp + 1 (assuming a one period review period), the average lead time demand, $\overline{D}_L = L\overline{D}$, and the standard deviation of lead time demand, $\sigma_L = \sqrt{L\sigma_D^2}$. In fact, the average lead time demand has to be forecasted as $\hat{D}_t^L = L\hat{D}_t$; $\hat{D}_t$ is the forecast of next period's demand, made in period t, and can be determined in different ways, e.g., moving average, exponential smoothing, long term average, or minimum expected mean squared error. We discuss these methods further in this paper. Of course, the standard deviation of lead time demand $\hat{\sigma}_L$ has to be estimated as well. In this paper, we assume however that σ_L is known and constant. This assumption simplifies the analysis, although the assumption is not realistic. Extending the analysis to include an estimated forecast error can be done easily (see Chen et al. 2000). Furthermore, the user can input a safety factor z to define the safety stock as $SS = z\sigma_L$ (Silver et al. 1998). However, any other safety stock value can be chosen. In this paper we will not elaborate on the determination of the safety stock. The amount of safety stock may be based on the economic stock-out probability (when shortage cost is known), or a predetermined customer service level or fill rate.

In order to evaluate the cost of the proposed policy, we input the following cost parameters: a holding cost C_h per unit per period, a backlog cost C_b per unit short, and a unit switching cost C_{sw} for changing the production level per period.

Next, the user can select a method to forecast customer demand. We distinguish five forecasting techniques: the mean demand forecast, the moving average forecast, the exponential smoothing forecast, the minimum mean squared error forecast and finally, demand signal processing. In the next section we discuss these forecasting techniques in detail. Once the forecast method is selected, the user can specify the parameters corresponding to the forecast method, respectively Tm, α and χ (to be discussed in the following sections).

2.2 Simulation

By clicking the "SIMULATE" button, a simulation of 500 periods is generated. The structure of the simulation table follows the sequence of events discussed earlier in the paper. We provide a screenshot of some periods in Figure 1. Every period, the incoming shipments from the upstream supplier are first received and placed in inventory. Assuming that the supplier has ample stock, these shipments correspond to the order placed $T_p + 1$ periods ago (T_p periods transportation delay and 1 period ordering delay). Next, a random customer demand is observed and either fulfilled (if enough on hand inventory available) or placed in backlog (corresponding to a negative net stock).

Insert Figure 1 About Here

The resulting net stock in period t is then equal to the net stock in the previous period, plus that period's receipt (equal to the order placed $T_p + 1$ periods ago), minus the observed customer demand. We also determine the number of items in the pipeline before an order is placed (WIP). The amount in the pipeline in the current period equals the pipeline amount of the previous period, plus the order placed at the end of the previous period, minus the order delivered this period. Hence we obtain

$$NS_t = NS_{t-1} + O_{t-(T_p+1)} - D_t , \tag{3}$$

$$WIP_t = WIP_{t-1} + O_{t-1} - O_{t-(T_p+1)}. \tag{4}$$

At the end of the period, a new order is placed to raise the inventory position (sum of net stock and inventory on order) to the order-up-to (OUT) level S_t :

$$O_t = S_t - (NS_t + WIP_t). \tag{5}$$

Note that we provide the *one-period ahead demand forecast* as well. We need this number to calculate the OUT level. In the next section we discuss in more detail how to obtain this demand forecast and the OUT level.

Finally, the costs per period are incurred. The inventory costs consist of a holding cost per unit in inventory (when net stock is positive) and a shortage cost per unit backlogged (negative net stock). The production switching costs are incurred for changing the level of production in a period. Assuming the production level is equal to the placed order quantity, the change in production is given by the difference in order quantity versus the previous period.

$$C_t^{INV} = \begin{cases} C_h \cdot NS_t & \text{if } NS_t \geq 0 \\ C_s \cdot (-NS_t) & \text{if } NS_t < 0 \end{cases} \tag{6}$$

$$C_t^{SW} = C_{sw} \cdot |O_t - O_{t-1}| \tag{7}$$

2.3 Output section

We define three types of performance measures of the simulation analysis: (1) the variance amplification ratios 'bullwhip effect' and 'net stock amplification', (2) the customer service measures 'customer service level' and 'fill rate' and (3) the average inventory and switching costs per period.

We define the bullwhip effect as follows:

$$\text{Bullwhip} = \frac{\text{Variance of orders}}{\text{Variance of demand}}.$$

A bullwhip measurement equal to one implies that the order variance is equal to the demand variance, or in other words, there is no variance amplification. A bullwhip measurement larger than one indicates that the bullwhip effect is present (*amplification*), whereas a bullwhip measurement smaller than one is referred to as a "smoothing" scenario, meaning that the orders are smoothed (less variable) compared to the demand pattern (*dampening*). When we know the variance of demand (which we assumed), we can verify our simulation results with the analytic results available in the literature. This is also the reason why we focus on a single echelon in our model. In a multi-echelon environment, the demand pattern of the upstream echelon is given by the order pattern of its downstream partner.

In general, however, we cannot determine the exact distribution of this order pattern, and therefore a comparative analysis with the analytic results available in the literature is hardly possible.

Our focus is not only on the bullwhip measure. In this paper we also check the variance of the net stock since this has a significant impact on customer service (the higher the variance of net stock, the more safety stock required). Therefore we measure the amplification of the inventory variance, NSAmp, as:

$$\text{NSAmp} = \frac{\text{Variance of net stock}}{\text{Variance of demand}}.$$

In case exact results for the bullwhip effect and net stock amplification are available in the literature, we provide them to compare with our simulated results.

The inventory and switching costs are related to these variance amplification measures. A high bullwhip measure implies a wildly fluctuating order pattern, meaning that the production level has to change frequently, resulting in a higher average production switching cost per period. An increased inventory variance results in higher holding and backlog costs, inflating the average inventory cost per period.

Finally, we provide the customer service level and fill rate resulting from the simulation analysis. The customer service level represents the probability that customer demand is met from stock, while the fill rate measures the proportion of demand that is immediately fulfilled from the inventory on hand.

Additionally we created some graphs to illustrate the bullwhip effect and the net stock amplification. By clicking on the "GRAPHS" button the user can observe the evolution of the simulated order pattern together with the observed demand pattern over time, and the simulated net stock evolution together with customer demand, both over a range of 50 and 500 periods.

3. IMPACT OF THE STANDARD ORDER-UP-TO POLICY ON THE BULLWHIP EFFECT

In the previous section we introduced the standard order-up-to policy: we place an order equal to the deficit between the OUT level and the inventory position (Eq. (1)). The OUT level S_t covers the forecasted average lead time demand and a safety stock:

$$S_t = \hat{D}_t^L + SS, \tag{8}$$

with $\hat{D}_t^L$ the forecasted demand over L periods and SS the safety stock (either equal to $z\sigma_L$ or set to an arbitrary value). There are two methods to calculate the forecasted demand over the lead time $\hat{D}_t^L$. The first is *one-period ahead forecasting* and is estimated by forecasting the demand of one period ahead and multiplying it by the lead time, i.e., $\hat{D}_t^L = L\hat{D}_t$, where $\hat{D}_t$ represents the forecast of next period's demand, made in period t. The second estimation method, called *lead time demand forecasting*, is calculated by taking the forecast of the sum of the demands over the lead time, $\hat{D}_t^L = \sum_{j=1}^{L} \hat{D}_{t+j}$, where $\hat{D}_{t+j}$ represents the j-period-ahead forecast, made in period t. In the first construction, the lead time is explicitly multiplicative, whereas in the second, the lead time is implicitly additive (see Kim et al. 2006). Unless stated otherwise, we assume one-period ahead forecasting in the remainder of this paper.

There are several ways to forecast demand. We will now review a number of forecasting techniques and illustrate their impact on the bullwhip effect by means of our spreadsheet models. We advise the reader to download the bullwhip explorer at http://www.econ.kuleuven.be/public/NDBAA78/BullwhipExplorer.xls; it makes it easier to follow the discussion below[1].

---

[1] If macros are disabled because the security level is set too high, the security level should be lowered to Medium with the Tools menu, Macro – Security submenu, before reopening the document.

The analytical results available in the literature are summarized in the Appendix (for both bullwhip and net stock amplification).

3.1 Mean demand forecasting

The simplest forecast method is mean demand forecasting. If the decision maker knows that the demand is IID, then it is quite clear that the best possible forecast of all future demands is simply the long-term average demand, $\overline{D}$. As a consequence, the forecasted lead time demand equals $\hat{D}_t^L = L\overline{D}$, and the OUT level S_t given by Eq. (8) remains constant over time, so that Eq. (1) becomes

$$O_t = S_t - (S_{t-1} - D_t) = D_t . \tag{9}$$

We simply place an order equal to the observed demand. That is why this policy is called the "chase sales policy". Consequently, in this setting, the variability of the replenishment orders is exactly the same as the variability of the original demand and the bullwhip effect does not exist.

By selecting in the spreadsheet model the "mean demand forecasting" technique, the user can observe how the generated orders are equal to the demand, with a bullwhip measure equal to one as a result. Although we do not discuss in this section the net stock amplification, it is worthwhile to check that number as well.

So why do we observe variance amplification in the real world? The answer is that decision makers do not know the demand (over the lead time) and consequently they forecast demand and constantly adjust the OUT levels. Suppose the demand is not characterized by an IID process, but rather a correlated or a non-stationary process, it is preferable to use the knowledge of the current demand to forecast next period's demand. Because of the fact that the true underlying distribution of demand is not directly observed (only the actual demand values are observed) many inventory theory researchers suggest the use of adaptive inventory control mechanisms (see e.g., Treharne and Sox, 2002). Unfortunately, these adjustments create bullwhip.

3.2 Demand signal processing

Lee et al (1997a) introduce the term "demand signal processing", which refers to the situation where decision makers use past demand information to update their demand forecast. As a result, the order-up-to level is not constant anymore, but it becomes *adaptive*. Suppose that the retailer experiences a surge of demand in one period. It will be interpreted as a signal of high future demand and the demand forecast will be adjusted and a larger order will be placed. Consequently the order-up-to level is adapted to

$$S_t = S_{t-1} + \chi(D_t - D_{t-1}),$$

resulting in the following order size:

$$O_t = O_{t-1} + \chi(D_t - D_{t-1}), \tag{10}$$

where χ is the *signaling factor*, a constant between zero and one. A value $\chi = 1$ implies that we fully adjust the order quantity by the increase (decrease) in demand from period to period.

Cachon and Terwiesh (2006) offer an excellent explanation for this ordering policy. An increase in demand could signal that demand has shifted, suggesting the product's actual expected demand is higher than previously thought. Then the retailer should increase his order quantity to cover additional future demand, otherwise he will quickly stock out. In other words, it is rational for a retailer to increase his order quantity when faced with an unusually high demand observation. These reactions by the retailer, however, contribute to the bullwhip effect. Suppose the retailer's high demand observation occurred merely due to random fluctuation. As a result, future demand will not be higher than expected even though the retailer reacted to this information by ordering more inventory. Hence, the retailer will need to reduce future orders so that the excess inventory just purchased can be drawn down. Ordering more than needed now and less than needed later implies the retailer's orders are more volatile than the retailer's demand, which is the bullwhip effect.

Suppose we select "demand signal processing" in our spreadsheet (the "Define a demand forecasting technique" window), then we immediately observe demand amplification. If we set $\chi = 1$, the bullwhip effect increases to a value around 5. If we anticipate to a lesser degree to the change of the demand, for example by setting $\chi = 0.2$, the bullwhip effect tempers to a value around 1.48. Observe that the switching costs also increase together with the bullwhip measure.

3.3 Moving average forecast

When the retailer does not know the true demand process, he can use simple methods to forecast demand, such as the moving average or exponential smoothing technique. This way future demand forecasts are continuously updated in face of new demand realizations. These estimates are then used to determine the order-up-to level (see Eq. (8)). Hence, adjusting the demand forecasts every period, the order-up-to level also becomes *adaptive*.

The *moving average* forecast (MA) takes the average of the observed demand in the previous periods. The one-period ahead forecast is given by

$$\hat{D}_t = \left(\sum_{i=0}^{T_m-1} D_{t-i} \right) / T_m , \qquad (11)$$

with Tm the number of (historical) periods used in the forecast. The forecast of the lead time demand is obtained by multiplying the one-period ahead forecast by the lead time L, $\hat{D}_t^L = L\hat{D}_t$, which determines the OUT level in Eq. (8).

By selecting the "moving average" forecasting technique in our spreadsheet models, we observe the impact of this forecast method on the order variability. Assuming an IID demand and a physical lead time of 2 periods, the bullwhip effect equals 3.63 for Tm = 4 (if one period corresponds to a week, then we use the demand data of the past 4 weeks or 1 month to compute the forecast). By using the data of 1 year or Tm=52, we obtain a much smaller bullwhip of 1.12 and we approach the chase sales policy. Indeed, the more data we use from the past, the closer our forecast will approach the average demand, and our results coincide with mean demand forecasting.

The spreadsheets also allow us to illustrate the effect of the lead times on the bullwhip effect. Doubling the physical lead time to 4 periods for example, the bullwhip measure increases to 6.63 with Tm = 4. The same results hold for an AR demand. We find that there is always bullwhip for all values of ρ and L. Clearly there is one exception that will result in no bullwhip (BW=1), namely when we set ρ = 0 and Tm=∞. In that case the AR(1) demand simplifies to the IID demand and the forecast equals the average demand, resulting in the chase sales policy.

3.4 Exponential smoothing forecast

The *exponential smoothing* (ES) forecast is an adaptive algorithm in which the one-period-ahead demand forecast is adjusted with a fraction of the forecasting error. Let α denote the smoothing factor, then the ES forecast of next period's demand can be written as

$$\hat{D}_t = \hat{D}_{t-1} + \alpha \left(D_t - \hat{D}_{t-1} \right). \tag{12}$$

Analogously to the moving average forecasting method, we multiply the one-period ahead forecast by the lead time L to obtain a measure of the lead time demand forecast.

We illustrate this forecasting method with our spreadsheets. When demand is IID and Tp=2, a smoothing factor α=0.4 generates a bullwhip measure of 5.20. We observe that an increase of α increases the bullwhip effect, since more weight is given to a single observation in the forecast. When α approaches zero (e.g. α = 0.001), we approximate the average demand as forecast. In that case the order-up-to level remains constant over time and hence there is no bullwhip effect (i.e. a bullwhip value of one). Similar to the MA forecast, we observe that an increase in the lead time results in a higher bullwhip measure.

3.5 Minimum Mean Squared Error forecast

Finally we consider the *minimum mean squared error* (MSE) forecasting method. With this forecasting technique, the demand forecast is derived in such a way that the forecast error is minimized. The MSE forecast for the demand in period $t + \tau$ equals the conditional expectation of $D_{t+\tau}$, given current and previous demand observations D_t, D_{t-1}, D_{t-2},... (Box and Jenkins 1976). Doing so, we exploit the underlying nature of the demand pattern to predict future demand. As a consequence it seems logic to explicitly forecast the τ-period-ahead demand to predict lead time demand, instead of simply multiplying the one-period-ahead forecast with the lead time (as in the MA and ES forecasting technique). Let $\hat{D}_{t+\tau}, \tau = 1,2,...,$ be the τ-period-ahead forecast of demand $D_{t+\tau}$ made in period t. Then,

$$\hat{D}_{t+1} = \overline{D} + \rho\left(D_t - \overline{D}\right), \tag{13}$$

$$\hat{D}_{t+\tau} = \overline{D} + \rho^{\tau}\left(D_t - \overline{D}\right). \tag{14}$$

The lead time demand forecast is obtained by plugging the τ-period-ahead forecast into the definition of lead time demand, $\hat{D}_t^L = \sum_{i=1}^{L} \hat{D}_{t+i}$. Hence, in contrast to the MA and ES forecast methods, we do not multiply the one-period ahead forecast with the lead time, but instead calculate the forecast of the demand over the lead time horizon L. The MSE forecast for lead time demand is then given by

$$\hat{D}_t^L = L\overline{D} + \frac{\rho - \rho^{L+1}}{1-\rho}\left(D_t - \overline{D}\right). \tag{15}$$

Clearly, the MSE forecasting scheme is optimal when demand is an AR(1) process, as it explicitly takes the correlative demand structure into account, which is not the case in the non-optimal MA and ES techniques. It assumes, however, that the underlying parameters of the demand process are known or that an infinite number of demand data is available to estimate these parameters accurately. When demand is IID ($\rho=0$), the above equations reveal that the MSE forecast reduces to mean demand forecasting. Note however that in the spreadsheet, only the one-period ahead forecast is given and not the lead time demand forecast.

We illustrate the impact of this forecasting method with our spreadsheets, and again assume Tp = 2. The results obtained are different from the previous results. When demand is negatively correlated, there is no bullwhip effect. When for instance $\rho = -0.5$, we obtain a bullwhip measure of 0.30, meaning that the order variability is dampened compared to the customer demand, instead of being amplified. We refer to Alwan et al. (2003) for a theoretical justification. When $\rho = 0.5$, we obtain a bullwhip measure of 2.64, so that the bullwhip effect is present for positively correlated demand. Note that when $\rho = 0$, the demand process is IID and the MSE forecast boils down to the mean demand forecast, resulting in a bullwhip measure of one. Furthermore, we again observe that increasing the lead time results in a higher bullwhip measure.

3.6 Insights

We have contrasted five different forecasting methods to replenish inventory with the standard order-up-to policy for both IID and AR(1) demand. The findings indicate that different forecasting methods lead to different bullwhip measures. The bullwhip measure also varies according to the lead time and demand process.

We conclude that, when we forecast a stationary demand based on its long term average and we keep the OUT level constant, there is no bullwhip effect. However, when we adapt the OUT level using a simple exponential smoothing, moving average or demand signal processing method, the standard order-up-to policy will always result in a bullwhip effect, independent of the demand process. The MSE forecasting technique is clearly the winner among the forecast methods, because it chases sales when demand is an IID process and it dampens the order variability when demand is negatively correlated. Moreover, it minimizes the variance of the forecasting error among all linear forecasting methods, and therefore it leads to the lowest inventory costs. Nevertheless, this forecast method requires an elaborate study to discover the parameters of the demand process.

We conclude that improper forecasting may have a devastating impact on the bullwhip effect. As a consequence, inventory and production switching costs may increase significantly. The spreadsheet application helps the decision maker to evaluate the impact of forecasting on the variability of the material flow. This observation puts forecasting in a totally different perspective.

4 IMPACT OF BULLWHIP REDUCTION ON CUSTOMER SERVICE

In the previous section we illustrated that the bullwhip effect may arise when using the standard order-up-to policy. In this section we introduce a generalized order-up-to policy that avoids variance amplification and succeeds in generating smooth ordering patterns, even when demand has to be forecasted. Smoothing models have a long tradition. The economic rationale of using smoothing replenishment (production) rules is quite obvious. A smoothing policy is justified when production (ordering) and inventory costs are convex (e.g. quadratic costs) or when there is a production switching cost. In such an environment it is preferable not to accept large deviations, instead some form of "averaging" is optimal.

In this section we introduce a generalized order-up-to replenishment rule. We discuss the trade-off between bullwhip and customer service, present some win-win solutions that arise for some specific demand patterns and discuss the use of the generalized OUT policy when demand is forecasted using the ES or MA forecast technique.

4.1 Generalized order-up-to policy

We present a generalized order-up-to policy with the intention of dampening the order variability or *smoothing* the order pattern. Consider the standard order-up-to policy. Substituting Eq. (8) into Eq. (1) we obtain

$$
\begin{aligned}
O_t &= \hat{D}_t^L + SS - IP_t = L\hat{D}_t + SS - IP_t \\
&= (T_p + 1)\hat{D}_t + SS - IP_t = \hat{D}_t + [T_p\hat{D}_t + SS - IP_t],
\end{aligned} \tag{16}
$$

where $T_p\hat{D}_t + SS$ can be seen as the *desired* inventory position DIP, which is the sum of the desired pipeline stock and desired net stock or safety stock. The difference between the desired and actual inventory position [DIP – IP_t] is denoted as the *inventory deficit*.

Introducing a proportional controller β for the inventory deficit, results in the following *generalized* order-up-to policy:

$$O_t = \hat{D}_t + \beta \cdot [DIP - IP_t], \qquad\qquad (17)$$

with $0 < \beta < 2$. Forrester (1961) refers to $1/\beta$ as the "adjustment time". When β < 1 he explicitly acknowledges that the deficit recovery should be spread out over time, whereas β > 1 implies an overreaction to the inventory deficit. Hence, when β < 1, the inventory deficit is only partially recovered during the next ordering period. This fractional adjustment is second nature to control engineers. It is the reason why the decision rule given by Eq. (17) may generate a "smooth" ordering pattern.

We developed a spreadsheet simulation of this generalized inventory policy[2]. The model is similar to the spreadsheet simulation of the standard OUT policy, but with a few important modifications. Additionally we input a value for the smoothing parameter β (since the control engineer literature prefers to use the inverse of β, namely Ti = 1/β, we also mention the Ti parameter in the input section). In Figure 2 we illustrate the impact on the order pattern when we choose a value β = 0.5, demand is IID and forecasted with its long term average. The fractional controller indeed has a dampened or "peak-shaving" impact on the order pattern; the resulting bullwhip measure equals 0.33.

Insert Figure 2 About Here

4.2 Trade-off between bullwhip and inventory variance

So far we have been concentrating on the variance of orders placed. This is, however, only one side of the coin. In developing a replenishment rule one has to consider the impact on the inventory variance as well, because that variance will have an immediate effect on customer service: the higher the variance, the more stock that will be needed to maintain customer service at the target level. We therefore measure the net stock amplification (NSAmp), which equals the ratio of the inventory variance over the demand variance.

Net stock variance (let alone variance amplification) is not a common supply chain measure, but we need it to calculate the fill rate, which is a popular customer service measure (see Disney et al. 2006).

Hence, we take into consideration the two following factors: on the one hand, the bullwhip effect which is related to the order variability and the switching costs, and on the other hand the the net stock amplification which is related to investment in inventories and the customer service.

Intuitively, we expect smooth ordering patterns will result in higher inventory fluctuations and consequently in a poorer fill rate, and this is indeed the case. Suppose we assume an IID demand, mean demand forecasting and Tp=2. A chase sales strategy with β=1 results in an NSAmp value of 3. Smoothing with β=0.5 reduces the bullwhip measure to 0.33, and equivalently decrease switching costs. On the other hand, it increases the NSamp measure to 3.33, together with an increase in inventory costs. We are able to smooth the order pattern, but pay the price of higher inventory fluctuations and more inventory costs.

These observations lead to a trade-off between bullwhip and customer service (as measured by net stock variance amplification). The question is to what extent production rates can be smoothed in order to minimize production adaptation costs, without adversely increasing our inventory costs too much (Disney and Towill 2003). Disney et al. (2004) show that when demand is IID and we forecast demand with its mean, then the sum of bullwhip and NSAmp is minimized at β = 0.618, which can be seen as "the best of both worlds" solution. This remarkable result is the "Golden Section", also known as the Golden Mean, Golden Ratio or Divine Proportion. By adding up the bullwhip effect metric and the net stock amplification metric, we assume that both factors are equally important. It is clear that in the real world companies apply weights to the bullwhip related costs and customer service related costs. In this case the shape of the total cost curve may be different and the optimal smoothing parameter may no longer be "golden".

---

[2] This model can be found in a second worksheet of the same file.

4.3 Win-win solutions for some demand patterns

We demonstrated that bullwhip can be reduced by ordering a fraction of the inventory deficit, rather than recovering the entire deficit in one time period. When demand is IID, order smoothing comes at a price: in order to guarantee the same fill rate, more investment in safety stock is required due to an increased inventory variance. Disney et al. (2006) show that it is possible to actually achieve bullwhip *and* inventory reduction together whilst maintaining customer service. This is a true win-win situation resulting from the generalized OUT policy. However, this cannot be achieved in all cases as it depends on the demand pattern.

Consider a stochastic demand pattern with auto regressive and moving average (ARMA) components of order one, i.e. ARMA (1,1), defined by:

$$D_t = \overline{D} + \rho\left(D_{t-1} - \overline{D}\right) + \varepsilon_t - (1-\delta)\varepsilon_{t-1}, \tag{18}$$

which is similar to the AR process, given by Eq. (2), but additionally contains a moving average component with δ the moving average coefficient, $0 \leq \delta \leq 2$ (Box and Jenkins 1976). Then, depending on the specific values of ρ and δ, inventory variance can be reduced by smoothing the demand signal ($\beta < 1$), so that bullwhip can be removed whilst reducing net stock variance (when compared to the standard OUT policy). In other cases, however, lower inventory variability is achieved by over-reacting to the ARMA signal (i.e., $\beta > 1$). If we then want to remove bullwhip in this situation, we are obliged to hold extra inventory.

These situations can be illustrated with the spreadsheets. For instance, suppose that ρ=0.5, δ=1.8 and we forecast demand with its long term average ("mean demand forecasting"). Then, a chase sales strategy (β=1) results in an NSAmp measure of 6.73. A value of β = 1.8 increases the bullwhip measure to 1.33, but decreases the NSAmp to 5.5 (observe that smoothing with β = 0.5 decreases the bullwhip to 0.66, but increases NSAmp to 9.13). Hence, in this case lower inventory variability is achieved with bullwhip. When we consider another example where demand is characterized by ρ=0.25 and δ=0.25, then a chase sales strategy (β=1) results in an NSAmp of 1.46.

Smoothing with $\beta = 0.5$ decreases the inventory variability to 1.15. Inventory variance is in this case reduced by smoothing the demand signal, which is a win-win solution. We refer to Disney et al. (2006) for a detailed analysis of potential win-win scenarios.

4.4 The generalized order-up-to policy with demand forecasting

We have to be cautious with the results described above. The smoothing rule described by Eq. (17) indeed provides the opportunity to dampen the variability in orders compared to the demand pattern. When an IID demand is forecasted with its long term average, Disney et al. (2006) show that the bullwhip measure is equal to $\beta/(2-\beta)$. Hence, for $0 < \beta < 1$ we generate a smooth replenishment pattern (dampening order variability) and for $1 < \beta < 2$ we create bullwhip (variance amplification). However, when the smoothing rule is applied and demand is forecasted with e.g. the moving average or exponential smoothing technique, the results are much more complicated. For instance, when demand is IID and forecasted with exponential smoothing and a smoothing parameter $\alpha = 0.5$, then a value $\beta = 0.5$ results in a bullwhip measure equal to 2.41. Hence the bullwhip effect is present, although the feedback parameter β is smaller than one. We should reduce β down to 0.2 in order to obtain a smooth order pattern with a bullwhip measure smaller than one when using this particular forecast method. In other words, improper use of forecasting techniques may destroy the smoothing effect of the generalized order-up-to policy.

The analytic formulas for bullwhip and NSAmp for the generalized OUT policy in the complete ARMA plane are very lengthy when we forecast using the ES technique. We refer to Disney et al. (2006) where these expressions are provided. For the purpose of this paper, we just mention that a feedback parameter $\beta < 1$ does not necessarily imply that the order variability is dampened when demand is forecasted with the MA or ES method. The decision maker can conduct several experiments with our spreadsheet model in order to obtain insights into this complicated issue.

4.5 Insights

When production is inflexible and significant costs are incurred by frequently switching production levels up and down, standard order-up-to policies with forecasting mechanisms may no longer be desirable. Because of the huge expenses, it may be important to avoid variance amplification or even to reduce variability of customer demand. Therefore we introduced the generalized order-up-to decision rule that avoids variance amplification and succeeds in generating smooth ordering patterns, even when demand has to be forecasted. The crucial difference with the standard order-up-to policies is that the inventory deficit is only fractionally taken into account. In developing this generalized replenishment rule, we have emphasized on two aspects: the order variability (as measured by the bullwhip effect), and the impact on the variance of the net stock (as measured by the net stock amplification). These performance measures are related to production switching and inventory costs.

The insights are clearest when we forecast demand with its long term average. When demand is an IID process, bullwhip reduction comes at a price. In order to guarantee the same fill rate, a larger safety stock is required. The "best of both worlds" solution, minimizing the sum of bullwhip and net stock amplification, is to set $\beta = 0.618$, the "Golden Ratio". For ARMA(1,1) demand patterns, it is possible to end up in four different scenarios when compared to the standard OUT policy: (1) *win-win*, we can remove bullwhip and reduce inventory; (2) *win-lose*, sometimes bullwhip can only be removed at the expense of holding extra inventory; (3) *lose-win*, sometimes bullwhip can be endured because it results in a policy that requires less inventory to be held; (4) *lose-lose*, sometimes excessive bullwhip and inventory may exist. These scenarios depend on the statistical properties of the demand process.

When demand is forecasted using the exponential smoothing or moving average method, the results are much more complex. In that case, a feedback parameter $\beta < 1$ does not necessarily imply that the order variability is dampened compared to the demand pattern. Using the spreadsheet application the decision maker can experiment with order smoothing and forecasting and as such, he/she can evaluate the impact of different replenishment strategies on the fluctuations in both the order and inventory pattern.

5. CONCLUSIONS

In this paper we present a spreadsheet application that can be used for educational purposes to gain a clear insight into the use or abuse of inventory control policies in relation to the bullwhip effect and customer service. It explores a series of replenishment policies and forecasting techniques under different demand processes, and illustrates how tuning the parameters of the replenishment policy and the forecasting technique result in the bullwhip effect. Moreover we demonstrate how bullwhip can be reduced, with an increase in inventory fluctuations and reduced customer service as a result. The spreadsheet models presented in this paper will guide the decision maker through a fairly complicated interplay between order fluctuations, inventory fluctuations and customer service in a variety of demand process scenarios and forecasting techniques.

ACKNOWLEDGEMENTS

REFERENCES

Alwan, L.C., Liu, J.J., Yao, D. (2003). Stochastic characterization of upstream demand processes in a supply chain. *IIE Transactions*, 35 (3), 207-219.

Box, G.E., Jenkins, G.M. (1976). *Time Series Analysis Forecasting and Control*. Holden-Day, San Francisco.

Cachon, G., Fisher, M. (1997). Campbell Soup's continuous product replenishment program: Evaluation and enhanced decision rules. *Production and Operation Management*, 6, 266-276.

Cachon, G. and Terwiesch, C. (2006). *Matching supply with demand: an introduction to operations management.* McGraw-Hill International Edition.

Chen, F., Samroengraja, R. (2000). The stationary Beer Game. *Production and Operations Management*, 9 (1), 19-30.

Chen, F., Ryan, J., Simchi-Levi, D. (2000). The impact of exponential smoothing forecast on the bullwhip effect. *Naval Research Logistics*, 47(4), 271-286.

Dejonckheere, J., Disney, S.M., Lambrecht, M.R., Towill, D.R. (2003). Measuring and avoiding the bullwhip effect: a control theoretic approach. *European Journal of Operational Research*, 147 (3), 567-590.

Disney, S.M. (2001). The production and inventory control problem in Vendor Managed Inventory supply chains, Ph.D. Thesis, Cardiff University, UK.

Disney, S.M., Towill D.R. (2003). On the bullwhip and inventory variance produced by an ordering policy. *Omega*, 31, 157-167.

Disney, S.M., Towill, D.R., Van De Velde, W. (2004). Variance amplification and the golden ratio in production and inventory control. *International Journal of Production Economics*, 90, 295-309.

Disney, S.M., Farasyn, I., Lambrecht, M., Towill, D.R., Van De Velde, W. (2006). Taming the bullwhip effect whilst watching customer service in a single supply chain echelon. *European Journal of Operational Research*, 173 (1), 151-172.

Forrester, J.W. (1958). Industrial Dynamics: A major breakthrough for decision makers. *Harvard Business Review* (July/August), 37-66.

Forrester, J.W. (1961). *Industrial Dynamics*. MIT Press, Cambridge MA.

Hammond, J. H. (1994). Barilla SpA (A). Harvard Business School case 6-694-046, Boston, MA.

Hosoda, T. (2005). The principles governing the dynamics of supply chains. PhD thesis, Cardiff University, UK.

Hoyt, D. (2001). Solectron: from contract manufacturer to global supply chain integrator. Stanford University Graduate School of Business, case GS-24, Stanford, CA.

Jacobs, F.R. (2000). Playing the Beer Distribution Game over the Internet. *Production and Operations Management*, 9 (1), 31-39.

Kim, J.G., Chatfield, D., Harrison, T.P., Hayya, J.C. (2006). Quantifying the bullwhip effect in a supply chain with stochastic lead time, *European Journal of Operational Research*, 173 (2), 617-636.

Kuper, A., Branvold, D. (2000). Innovation diffusion at Hewlett Packard, in M.E. Johnson and D.F. Pyke, eds: *Supply Chain Management: Innovations for Education*. Production and Operations Management Society, Miami, FL, 205-218.

Lee, H.L., Padmanabhan, V., Whang, S. (1997a). The bullwhip effect in supply chains. *Sloan Management Review*, 38 (3), 93-102.

Lee, H.L., Padmanabhan, V., Whang, S. (1997b). Information distortion in a supply chain: the bullwhip effect. *Management science*, 43(4), 543-558.

Lee, H.L., Padmanabhan, V., Whang, S. (2004). Comments on "Information distortion in a supply chain: the bullwhip effect". *Management Science*, 50 (12), 1887-1893.

Machuca, J.A.D. and Barajas, R.P. (1997). A computerized network version of the beer game via the internet. *Systems Dynamics Review*, 13 (4), 323-340.

Nahmias, S. (1997). *Production and Operation Analysis*. McGraw-Hill, 3rd edition.

Naim, M.M. and Towill, D.R. (1995). What's in the pipeline? *Proceedings of the 2nd International Symposium on Logistics*, pp. 135-142.

Peleg, B. (2003), STMicroelectronics e-chain optimization project: Achieving streamlined operations through collaborative forecasting and inventory management. Stanford Global Supply Chain Management Forum case, Stanford University, Stanford, CA.

Silver, E.A., Pyke, D.F., Peterson, R. (1998). *Inventory Management and Production Planning and Scheduling*. John Wiley & Sons, New York, 3rd edition.

Simchi-Levi, D., Kaminsky, P. and Simchi-Levi, E. (1998). *Designing and Managing the Supply Chain*. Irwin/McGraw-Hill, New York.

Sterman, J. (1989). Modeling managerial behaviour: misperceptions of feedback in a dynamic decision making experiment. *Management Science*, 35 (3), 321-339.

Treharne, T., Sox, C. (2002). Adaptive inventory control for non-stationary demand and partial information. *Management Science*, 48, 607-624.

Watson, N. and Zheng, Y.S. (2002). Adverse effects of over-reaction to demand changes and improper forecasting. Working paper, Harvard Business School.

Zhang, X. (2004). The impact of forecasting methods on the bullwhip effect. *International Journal of Production Economics*, 88, 15-27.

Zipkin, P.H. (2000), *Foundations of Inventory Management,* McGraw-Hill, New York.

APPENDIX

Table 1: Exact formulas for the standard order-up-to policy.

Table 2: Exact formulas for the generalized order-up-to policy.

Table 1: Exact formulas for the standard order-up-to policy. Sources: (1) Dejonckheere et al. 2003, (2) Chen et al. 2000, (3) Hosoda 2005, (4) Zhang 2004

	IID		AR(1)	
	Bullwhip	**NSAmp**	**Bullwhip**	**NSAmp**
Mean demand forecasting	1	$1+Tp$ [3]	1	$1+Tp+\dfrac{2\rho(Tp-Tp\rho+\rho(\rho^{Tp}-1))}{(\rho-1)^2}$ [3]
Moving Average	$1+\dfrac{2L}{Tm}+\dfrac{2L^2}{Tm^2}$ [1,2]	$\dfrac{L(L+Tm)}{Tm}$ [3]	$1+\left(\dfrac{2L}{Tm}+\dfrac{2L^2}{Tm^2}\right)(1-\rho^{Tm})$ [2,4]	$\dfrac{\left(\begin{array}{l}L^2(Tm(1-\rho^2)-2(\rho-\rho^{Tm+1}))+\\ LTm(Tm(1-\rho^2)-2(\rho-\rho^{L+1})(1-\rho^{Tm}))+\\ 2Tm^2(\rho^{L+1}-\rho)\end{array}\right)}{Tm^2(\rho-1)^2}$ [3]
Exponential Smoothing	$1+2L\alpha+\dfrac{2L^2\alpha^2}{2-\alpha}$ [1,2]	$\dfrac{L(\alpha-2)-L^2\alpha}{\alpha-2}$ [3]	$1+\left(2L\alpha+\dfrac{2L^2\alpha^2}{2-\alpha}\right)\left(\dfrac{1-\rho}{1-(1-\alpha)\rho}\right)$ [2,4]	$\dfrac{\left(\begin{array}{l}L^2\alpha(\rho-1)^2(\alpha\rho-\rho-1)+L(\alpha-2)(1-\rho)\\ (1-\rho^2+\alpha\rho(2\rho^L+\rho-1))+\\ 2(\alpha-2)\rho(\alpha\rho-\rho+1)(\rho^L-1)\end{array}\right)}{(\alpha-2)(\rho-1)^2(\alpha\rho-\rho+1)}$ [3]
Demand signal processing	$1+2\chi(1+\chi)$ [1]	$1+Tp+\chi^2$	/	/
MMSE	1	$1+Tp$	$1+\dfrac{2\rho(1-\rho^L)(1-\rho^{L+1})}{(1-\rho)}$ [3,4]	$\dfrac{L(1-\rho^2)+\rho(1-\rho^L)(\rho^{L+1}-\rho-2)}{(1-\rho)^2}$ [3]

Table 2: Exact formulas for the generalized order-up-to policy. Source: Disney et al. 2006

	IID	
	Bullwhip	**NSAmp**
Mean demand forecasting	$$\dfrac{\beta}{2-\beta}$$	$$1+Tp+\dfrac{(1-\beta)^2}{(2-\beta)\beta}$$
Exponential Smoothing	$$\dfrac{-2\beta^2+\alpha\beta(-6+(3-4Tp)\beta)-\alpha^2(2+\beta(-3+\beta+2Tp(2+(-1+Tp)\beta)))}{(\alpha-2)(\alpha(\beta-1)-\beta)(\beta-2)}$$	$$-\dfrac{2+Tp(2+\alpha(\beta-1)-\beta)(2\beta+\alpha(2+(Tp-1)\beta))}{(\alpha-2)(\alpha(\beta-1)-\beta)(\beta-2)}$$
	AR(1)	
	Bullwhip	**NSAmp**
Mean demand forecasting	$$\dfrac{\beta}{2-\beta}\dfrac{1-(\beta-1)\rho}{1+(\beta-1)\rho}$$	$$\dfrac{\beta\left(-\dfrac{(-1+Tp(\beta-2)\beta(-1+(\beta-1)\rho)}{(\beta-2)\beta^2}+\dfrac{2\rho(Tp-Tp\rho+\rho(\rho^{Tp}-1))}{(\rho-1)^2}\right)}{1+(\beta-1)\rho}$$
Exponential Smoothing	See Disney et al. 2006	See Disney et al. 2006

FIGURE 1

Spreadsheet example of a standard OUT policy with Tp=2

period	receive	demand	NS	WIP	demand forecast	OUT-level	order	inventory costs	switching costs
10	89	109	36	187	104,00	331,50	109	18,00	44,00
11	100	100	36	196	104,50	333,00	101	18,00	16,00
12	87	102	21	210	101,00	322,50	92	10,50	18,00
13	109	105	25	193	103,50	330,00	112	12,50	40,00
14	101	105	21	204	105,00	334,50	110	10,50	4,00
15	92	111	2	222	108,00	343,50	120	1,00	20,00
16	112	107	7	230	109,00	346,50	110	3,50	20,00

FIGURE 2

Generated order pattern when β = 0.5