Managing bullwhip induced risks in supply chains

Denis R Towill and Stephen M Disney
Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University,
Aberconway Building, Colum Drive, Cardiff, CF10 3EU, Wales, UK.

Summary

We discuss the exploitation of a well established replenishment rule, the order-up-to policy to control the supply chain risk resulting from the bullwhip effect. To avoid this risk within supply chains, two specific recommendations are highly important. Firstly we provide strong evidence that availability of high fidelity up-to-date delivery lead time information is essential. This inevitably requires transparency and trust between various “players” within the supply chain. For many practical reasons this is easier to recommend than to actually achieve. Secondly, such reliable data needs to be automatically fed into the replenishment system to enable it to function in an adaptive mode. This step offers a reasonable guarantee of well controlled performance. It is also advisable that at the design stage, replenishment rule parameters should be set so that “conservative” operation is achieved. The new theoretical advances described in the paper are validated via simulation models of the delivery process.

Key words: Bullwhip, Inventory variance, lead-times, lead-time estimation errors, stability, APIOBPCS

1. Introduction

In this paper we study how supply chain risks may affect company profits by increasing lead times and inducing bullwhip problems. To achieve this goal we first review the concept of bullwhip, and its distinguished predecessor, demand amplification. The latter phenomenon was superbly demonstrated via simulation by Jay Forrester, (1958) “the Prophet of Unintended Consequences” (Fisher, 2005). Nearly four decades later Lee et al, (1997) coined the “bullwhip” phrase to describe similar events. The latter terminology has since become extremely popular. However, this phenomenon is particularly well known to economists (Mitchell, 1923), and Chandler (1990) has quoted an example occurring in the Proctor and Gamble supply
chain as far back as 1919. Such “boom and bust” behaviour can be extremely costly, as the seminal study by Metters (1997) has established. He shows it can easily add 30% on-costs at a single stage in the supply chain. Since the “domino” effect is very strong in such a scenario it follows that the risk to profit is high.

To minimise this bullwhip risk it is essential to realise the practical limitations of real-world supply chain operations. Firstly lead times may occasionally be relatively small but they are non-zero. However it is also important to emphasise that for smooth supply chain operations they need to be both reasonably short and consistent and known. Secondly, demand cannot be perfectly forecasted as has been adequately demonstrated, even in the relatively stable world of the sale of Japanese motor vehicles (Mollick, 2004). Hence the practical need for compressing delivery lead times (Stalk and Hout, 1990) which in turn drastically reduce forecasting errors (Watson, 1993). Such a goal is thereby very simply summarised, but much more difficult to achieve in the extended enterprise. This is because there may be many echelons present between raw material supplier and end customers. Hence there will be conflicting vested interests targeted to suit individual businesses. So, regrettably, not all “players” are necessarily aiming for the goal of end customer satisfaction via establishing and maintaining a seamless delivery system (Towill, 1997).

This practical scenario is exacerbated because individual decision makers such as a production scheduler or store manager may be responsible for overseeing the processing of orders for several thousand stock keeping units (SKU’s) every week, or may be, even every day. Hence the need for robust autonomic Decision Support Systems (DSS) becomes paramount. Typically such a decision maker would expect to process, say, 95% of such orders algorithmically. The remainder may require further detailed consideration on such grounds as criticality to business or unusual supplier scenarios. The last thing required in such a high pressure activity is the breakdown of such an algorithmic routine. The system is then in severe danger of overload as the apparently “routine” SKU’s become a real-time problem. This often leads to a panic reaction and consequently poor decision making.

In a previous paper we have provided suitable DSS guidelines to select and parameterise these replenishment rules so as to avoid bullwhip occurring (Disney and

Towill, 2006a). These recommendations worked extremely well under conditions where actual lead-times and perceived lead-times were matched. In this paper we extend the theory to predict system response where a mismatch occurs. This could be by an accident within a factory, by a machine or transport problem or due to “players” withholding such information. We show that such a situation considerably increases the possibility of bullwhip occurring. Hence the risk of profits being lost due to the bullwhip phenomenon is now very much greater. By extending our knowledge of system stability we demonstrate how such risk is kept under control.

1.1. Demand amplification in the real world

Since the seminal simulation based research of Forrester (1958), there has been considerable activity in this field as typified by Christenson and Brogan, (1971), Roberts (1981), Wikner et al (1991), van Ackere et al, (1993) and Sterman, (2000). But what actually happens in the real world of business and commerce? Figure 1 is a typical example of demand amplification occurring in a supermarket (Jones et al, 1997). Retail sales appear by eye to be reasonably constant over the time period observed. However there is clearly some variation with the extreme swings being approximately ± 25% of average sales. But variance of the orders placed on the suppliers is much larger. Now the extreme swings are roughly ± 50% of average sales. In other words, based on deviations about the observed average, the bullwhip evident in this simple scenario is about 2:1.

However the volatile realities for transportation in the supply chain are much more extreme. If the supplier delivers exactly what is ordered by the retailer, his requirements for trucks will vary by even more than 2:1. For example, the capacity needed could drop to as low as, say, three trucks but can rise as high as six trucks. In contrast, if the supplier was required to replenish sales as they actually occur only three to five trucks are required; although four trucks would be adequate for most deliveries. This variation as seen in the UK retail sector example appears largely random, maybe with some evidence of “bunching” orders to the supplier.
In contrast the “spikiness” of the pattern of behaviour evident in Figure 2 is clearly due to an “event”. The cause is a big annual marketing push resulting in the self-evident peak and subsequent to the discounted sales promotion (Fisher et al, 1997). So although the actual US consumption of Campbell’s Soups is clearly seasonal, with a very strong winter component roughly centred at Christmas, the orders become highly distorted. Thus the shipments that are needed to ensure delivery to retail outlets has peaked by twice the maximum actual consumption and as much as four times the average sales rate. Even more spectacularly, the shipments peak is nineteen times the low values as recorded between weeks 34 to 41. This is bound to result in a significant amount of unnecessary costs at several echelons. The consequence of this discounted sales promotion may be an apparent short term increase in market share. But this is achieved at the expense of necessitating a highly distorted pipeline impacting on every business in the supply chain. So it is not surprising that bullwhip scenarios get rapidly worse not better as we look at the complete picture.
A simple conceptual model of a clothing supply chain is shown in Figure 3. Materials and products flow downstream and orders flow upstream from echelon to echelon. Each echelon (retailer: garment maker: fabric maker: and yarn maker) is subject to delays in process. It also needs to make a decision on passing orders to “their” vendor in the light of such information as “orders” received from “their” customer, work-in-progress and stock levels, Stalk and Hout (1990). Hence it can take many months for a complete cycle of orders to pass up the chain to the original source and for the corresponding products to flow down to the marketplace especially if overseas outsourcing is used (Lowson, 2001). So, by the time the goods actually arrive in the shops the customer probably wants something quite different to what was originally forecast (Mansell, 2002). Bullwhip is an almost inevitable consequence of such a long and protracted chain, with every “player” double-guessing on what is really required.

In Figure 3 we have arbitrarily shown the bullwhip increasing by a factor of 2:1 as orders pass across each interface. This is actually modest compared to these results reported by Forrester (1958) or van Ackere et al (1993). Since bullwhip tends to be multiplicative, this means that across the complete supply chain the demand amplification is likely to be 8:1. This is regrettably typical of real-world supply
chains. Again, this is not a new phenomenon, as the excellent early description of bullwhip occurring in the retail sector testifies (Mitchell, 1923).

![Diagram of a clothing supply chain]

**Figure 3. Generation of bullwhip in a clothing supply chain**
(Source: Stalk and Hout, 1990)

A relatively crude (but extremely valuable) set of bullwhip estimates within a European confectionary supply chain have been published by Holmström (1997). These are shown in Table 1. The demand amplification measure used here is the statistical metric of the variance estimated at each stage. Since under certain conditions variances multiply it can be seen that the total effect throughout the chain is quite horrendous. But in this real-world supply chain the production scheduler manifestly reacts differently to orders for these two sample products. Either due to market intelligence or just gut feeling, he realises that the high volume product is a relatively stable seller. So he sensibly adopts a level scheduling strategy on orders arriving at the factory from the wholesalers. Such a decision may well help reduce acquisition on-costs. But, of course, it may incur some hidden costs, such as those due to lost business resulting from stock-outs.
Figure 4 shows the conceptual problem tackled in this paper. For finite demand variance and a finite lead-time there will be either order variance, and/or inventory variance. Both sources contribute on-costs in the delivery system, and hence drive down profit. Now it is possible to implement an appropriate Decision Support System (DSS) which will minimise such costs where the lead-time is known (and fixed), and the demand statistics are available. Unfortunately it is easy to induce bullwhip into such a situation via a poor choice of DSS algorithm and/or unsuitable parameter settings. The upside of this argument is clear, however. With the requisite information available, it is straightforward to achieve low bullwhip and low on-costs (from this source) by knowledgeable DSS selection and good timing of the algorithm. Furthermore it may also be reasonably robust to some changes in operating conditions (Disney and Towill, 2006a).
Regrettably, in the real-world the operating conditions will rarely remain fixed. What should the supplier company then do? There are two lines of action recommended to reduce the risk of added bullwhip costs impinging negatively upon profits. The first is *action* to minimise change. Hence on the demand side there should be interaction with customers so as to avoid excessive peaks and troughs in their orders placed. In other words wherever possible the self-smoothing demand is an integral part of company strategy (as is indeed the case with the Toyota Production System, Suzaki 1987). But there can also be appropriate action taken on the incoming value streams from *our* vendors. Hence when US stevedores industrial action interfered with component supplies, Dell despatched special task forces to both local and overseas docksides. This was to ensure that wherever and whenever possible, loading and unloading of Dell components and sub-assemblies were highly prioritised (Holzner, 2006). Additionally, if necessary, aircraft were chartered to help protect the supply system against any such disruptions. At the first sign of trouble they would be reserved for Dell use.

Other possible activities which maybe similarly classified as extended action to reduce the risk of on-costs include pro-active searches for alternative vendors and re-
distribution of supplier business between the existing ones (Nishiguchi and Beaudet, 1999). The latter initiatives were indeed taken by Toyota when their sole supplier of a key brake sub-assembly business was temporarily incapacitated because of a major factory fire. However this pathway to risk reduction via interaction with customers and suppliers in the manner described is a separate issue and is not further discussed in this paper. Nor is the associated principle of “substitution” considered further. In the latter venture there is a strong endeavour to replace a sub-assembly (with disrupted deliveries) with the nearest equivalent (which may have ample stock available). This concept could even be extended to the salesman persuading the customer that there are advantages in his purchasing an old model (in stock) as being superior for purpose to the newly advertised model. Unbeknown to the client the updated version may well have major components immobilised elsewhere around the world because of a stevedore’s strike (Papadakis, 2003).

Interesting and valuable that such actions are, attention in this paper is concentrated on the other complementary pathway for reducing bullwhip induced on-costs risk. This is the information pathway, in which lead-time updates are automatically exploited within the DSS. Davenport (2006) looks forward to the full implementation of the IT age in business. He paints a highly relevant scenario in which companies compete on the basis of their analytic capability. In such a state of operation the type of information usage we propose should become endemic. Consequently this paper concentrates on the important principle of reducing risk or excessive bullwhip on-costs via close monitoring of actual lead-times.

These estimates are fed into an adaptive update of the DSS within the replenishment feedback loop (Towill et al, 1997). The latter system was developed pragmatically. Herein we test the principle widely via a range of operating scenarios based on typical industrial data. This amply demonstrates the improvements made possible via lead-time adaptive controls. Additionally we provide a novel stability theory which considerably aids understanding of “best” parameter settings. Hence the action information pathway is shown to be readily implemented. It is consistent with, and complementary to any action pathway which the extended enterprise is engineering.
1.3. Contribution of this paper
In supply chain operations there is an inevitable trade-off between the “physical” (that is, acquisition) costs, and the “market mediation” or implicit costs. The latter can be well hidden from view at times of stable market conditions. However there are many times some which are unrecognisable when it will dominate the supply chain Product Delivery Process (PDP) costs. The Fisher et al, (1997) equation is:

\[
\text{Total supply chain PDP costs} = \left( \text{Physical PDP costs} \right) + \left( \text{Market mediation PDP costs} \right)
\]

(1)

where the physical PDP costs include all production, distribution, storage and similar costs, and the market mediation PDP costs include all obsolescence stock-out and similar costs.

For the purposes of this paper we regard the supply chain risk to the PDP as essentially dominated by the bullwhip effect. For such a context the Sterman, (2000) “boom-and-burst” descriptor is extremely apt. During the boom phase the risk of incurring excess costs in Equation 1 includes all the extra costs paid for rapid acquisition of more materials, overtime premiums and learning curve effects, and frequently compounded by the impact of lower quality upping the reject rate. Hence boom induced risks of excessive costs are primarily impacting the physical part of Equation 1, plus the loss of profit due to stock-outs.

In contrast when the downstream “bust” occurs the businesses throughout the chain are stuck with unwanted stock. This is lying around at many levels in the supply chain. Hence excess holding costs are immediately incurred. But the risk to profit does not end there, since two further negative factors come into play. The first is fairly obvious, since stock is written down, and disposed of in the inevitable retail “sale”. Prices (and hence profit margins) are slashed just to generate cash and to make room for the next delivery (which may already be on its way). But there is still the very subtle factor of “cannibalisation”. In this case the canny customer will buy two items at the reduced price now. One is to use immediately, but the second is an
advance purchase of a substitute for next year’s model. Hence the chance of a full price sale next year is also foregone.

As we have argued previously, at any particular echelon within the supply chain it is clearly a good idea to avoid generating unnecessary bullwhip. We have previously detailed a suitable algorithm which achieves this goal under carefully defined conditions (Disney and Towill, 2006a). We now extend that knowledge by considering the deleterious effect due to mismatched lead-times when setting target inventory and work-in-progress (WIP) levels. We show the extent to which this impacts on supply chain volatility and develop a new theory to establish stability limits as a function of this mismatch. Our results greatly strengthen previous arguments for enabling transparency of rapidly circulated high fidelity information within the supply chains. This will considerably reduce risks to profit margins, because the likely cost of implementation will be swamped by avoidable bullwhip on-costs.

2. Our supply chain replenishment rule
The particular replenishment rule advocated and exploited in this paper is a special case of the automatic pipeline inventory and order based production system (APIOBPCS). It is designed to control delivery when the actual lead-time is $T_p$. This rule readily expressed in words as,

Let the replenishment orders be equal to the sum of an exponentially smoothed demand (averaged over $T_a$ time periods), plus a fraction ($1/T_i$) of the inventory difference between target stock and actual stock, plus a fraction ($1/T_w$) of the difference between target “orders placed but not yet received” and actual “orders placed but not yet received” (John et al., 1994).

John et al (1994) also highlights the importance of utilising the “best” lead-time estimate, $T_p$, currently available in setting the target “orders placed but not yet received” (or WIP level) if inventory drift is to be reduced. As we shall see later, the practical benefit of so doing is in fact much greater than just solving this previously identified problem.
The APIOBPCS model encapsulates the general principles for replenishment rules as advocated by Popplewell and Bonney (1987). In particular it gives due prominence to the importance of including pipeline (WIP) feedback in replenishment decision making, a factor further emphasised by Bonney (1990). Of course the APIOBPCS principle is not a new concept. It is empirically well established in industry (Coyle, 1977), and has the additional advantage of reasonably describing performance data from 2000 Beer Game “plays” as elegantly modelled by Sterman (1989). Furthermore the particular variant known as the “to-make model” has successfully if rather pragmatically controlled 6,000 multi-product pipelines in the UK orthopaedic components industry. This was via exploitation of empirically derived parameter settings (Cheema et al, 1989).

Deserved customer service levels can be ensured with the APIOBPCS model by the use of a target stock level. This can be set arbitrarily and does not affect the stability or variance ratios between the demand variance and the order (or inventory) variance in a linear system (an assumption we make herein). Thus in theory, with a high enough target stock level, any desired availability or fill-rate (or any other measure) can be achieved. Indeed, by exploiting the classic “newsboy” principle, the target stock level can be set to the critical fractile that ensures the optimum economic stock-out probability is achieved, Disney et al (2006b).
A very simple, but operationally profound, and highly recommended modification to APIOBPCS is undertaken by making $T_w = T_i$. This simplified model is designated DE-APIOBPCS as this streamlining was first advocated by Deziel and Eilon (1967) researching within an OR context. A simple but powerful demonstration of the importance of the DE settings of $(T_w = T_i)$ is available in the analysis by Disney and Towill (2003) and it generally results in a very good dynamic response. Such a modification also has the great benefit of hugely simplifying the subsequent mathematical analysis. Typically the estimation of bullwhip is much more tractable and we are able to generate the unit bullwhip contours shown in Figure 5. By setting the replenishment rule parameters to the values shown on the graph then for idealised conditions and random demand, bullwhip is avoided if $T_p = T_p$.

In theory Figure 5 provides the user with helpful guidelines based on mathematical analysis. This contrasts with previous pragmatic approaches such as Towill et al. (1997) to parameter settings in which bullwhip levels were inferred rather than calculated. The relevant bullwhip and inventory calculations for system synthesis
may be undertaken from the Appendices and are listed in Table 2. The full implications of these results will be discussed in the next section.

| Design | Replenishment rule settings (based on expected lead time) | Predicted performance | | | |
|---|---|---|---|---|
| 1 | 1/(1+T_p) | Bullwhip Inventory variance | Bullwhip Inventory variance | Bullwhip Inventory variance |
| 2 | 1/(1+T_p) | 2.205 | 1.352 | 4.781 | 1.028 | 4.946 |
| 3 | 1/(1+T_p) | 2.392 | 1.586 | 5.432 | 1.148 | 6.229 |
| 4 | 1/(1+T_p) | 2.541 | 1.727 | 5.915 | 1.747 | 9.609 |
| 5 | 1/(1+T_p) | 2.417 | 1.612 | 5.515 | 2.848 | 13.498 |

Table 2. “Steady State” performance for the APIOBPCS replenishment rule
(Source: Authors)

3. Impact of mismatched lead-times
In a previous paper we had benchmarked bullwhip via mathematical analysis for the design case where T_p=T_p (Disney and Towill, 2006a). Hence the first two sections of Table 2 result for (T_p=T_p=1; and for T_p=T_p=3) are already established in the literature. In other words a reasonable degree of system robustness is self-evident. For example, Design 1 has unity bullwhip for the expected value of T_p=1, but which has a modest increase up to 1.352 if T_p increases to 3 (provided this is known and used to adjust T_p accordingly). Design 5 is also comparable in nominal performance, so if the appropriate information is available and used, no unexpected problems arise. This broad equivalence is confirmed in the simulated responses shown in Figure 6.
But in reality the situation is actually much worse than suggested by the “steady state” theory of Table 2. This is because as shown in Figure 7, Design 5 exhibits a very lightly damped response mode which should always be a cause for concern. As we show in the Appendices, for a stable system our theory has now been advanced to estimate bullwhip and inventory variance for the more challenging situation in which $T_p \neq T_p$. Thus, the right hand section of Table 2 is $T_p=3$, $T_p=1$ is completely new and novel. It is now immediately apparent from these results that this large lead time...
mismatch has destroyed the previously apparent congruence between Design 1 and Design 5. Whereas the former system has coped with the lead-time “ignorance” very well, the latter is now quite volatile. As a consequence risk of profit reduction is much greater. In other words, for such a system the “tipping point” (Gladwell, 2000) for a potentially disastrous instability may be close at hand.

(a) Design 1 when \( T_p = 1 \) and \( T_p = 3 \) throughout the simulation

(b) Design 5 when \( T_p = 1 \) and \( T_p = 3 \) throughout the simulation

**Figure 7. Different behaviours with mismatched lead times**  
(Source: Authors)
4. A solution to this problem: Effective adaptive control

We now show that adaptive control of the replenishment rule does indeed provide a much better response and hence reduced risk of bullwhip lowering profitability. Figure 8 shows schematically how this is enabled by building a current lead time estimator into the WIP feedback path. Hence, in the adaptive system the WIP target is automatically adjusted in line with the observed actual value of $T_p$. As Towill et al (1997) has posited this updating process can itself be a variable and some form of exponential smoothing may be a useful option. However for the purpose of illustration herein, we shall assume the adaptive loop operates instantaneously when the new information is received. Smoothing could, of course, be helpful if the data source used for updating the lead time estimate is “noisy”.

![Figure 8. Adaptive target WIP loop added to the replenishment system](image)

(Adapted from Towill et al, 1997)

Figure 9 shows the simulated response of the delivery system with and without the proposed adaption. The differences in performance are quite startling. Whereas the non-adaptive system is extremely volatile, the adaptive scheme results in a well damped response. The former will certainly induce significant additional costs (and hence reduce profits) due to the completely system induced cyclic behaviour about the
trend line. Furthermore such on-costs are unlikely to be restricted to this particular supply chain echelon. There could additionally be significant knock-on effects at all upstream stages as well. As Burbidge (1984) remarked, in his vast experience if there is any chance or such a domino effect happening, then it almost certainly will.

(a) **Non-adaptive controls.** Design 5 when $T_p = 1$ and $T_p=1$ until time 50, then $T_p = 1$ and $T_p=3$. There are 2 period of zero receipts (time periods 50 and 51)

(b) **Adaptive controls.** Design 5 when $T_p = 1$ and $T_p=1$ until time 50, then $T_p = 3$ and $T_p=3$. There are 2 period of zero receipts (time periods 50 and 51).

**Figure 9.** Improved dynamic response via adaptive control implementation
(Source: Authors)
4.1. Extended stability analysis

Disney and Towill, (2005) have previously shown the stability condition for the particular case where actual lead time and expected lead time are equal turns out to be beautifully simple. The value for $T_i$ (when $T_w=T_i$) just has to be greater than 0.5, and this parameter value is independent of $T_p$ and of course $T_p'$. The actual response will vary according to the value of $T_i$ selected, as the bullwhip and inventory variances listed in Table 2 confirm. However the earlier paper left the issue of any problems arising from $T_p \neq T_p'$ to be resolved by the use of a suitable simulation model to enable answers to “what if” questions. In other words the results previously shown in Figure 7 are typical of what might be expected.

![Stable system](image1)

**Stable system**
$T_p=1, T_B=1, T_a=\infty, T_i=1$

![Unstable system](image2)

**Unstable system**
$T_p=2, T_B=1, T_a=\infty, T_i=1$

![Critically stable system](image3)

**Critically stable system**
$T_p=3, T_B=1, T_a=\infty, T_i=1$

Figure 10. Confirmation of the stability boundary due to mismatch in lead times
(Source: Authors)
But this particular output has caused us to investigate the problem in even greater depth as a prelude to reducing bullwhip induced risk. Hence our theory has now been greatly extended via Appendix I (for stability analysis) and via Appendix II (for subsequent estimation of bullwhip and inventory variance) for a stable system. The concerned mathematics is rather specialised and tedious. However the dramatic output of Figure A1 is of interest in two important respects. Firstly there is no longer a simple stability relationship between $T_i$ and $T_p$ as it clearly depends on $T_p$. Secondly the “safe zone” for system operation is now a peculiar shape in the $T_i$ versus $T_p$ plane. This is due to the contours for a given value of $T_p$ being discontinuous in slope at more than one point in the plane.

How these results affect parameter settings and operating scenarios will be discussed in detail in the next section. Here we are concerned with cross-checking and hence validating the theoretical results shown in Appendix A. To do this we have simulated three different systems, all with the expected lead time $T_p = 1$ unit, reacting to a unit impulse in demand. The control parameters have been arbitrarily kept at $T_u = \infty$, $T_i = T_w = 1$. However $T_p$ takes on the values 1, 2, and 3 in turn. The simulation output to an arbitrary input are shown in Fig. 9. As expected, when $T_p = 1$ we have a well behaved stable response. But if the actual lead time increases to $T_p = 2$ the system becomes dramatically unstable (note the vertical scale has been substantially changed to make the point that this response is widely fluctuating). Finally, increasing $T_p = 3$ results in a continuously oscillating system. This is critically stable, but of course is still risking reduced profit via unnecessary volatility when the demand is actually quiescent.

5. Procedures for reducing bullwhip induced risk

The kind of situation reviewed in this paper is not uncommon in industrial operations. For example, execution at a UK manufacturer with a value stream audited via the “Quick Scan” procedure (Naim et al 2002) recognized the importance of minimising lead-times. It therefore implemented a performance related pay system for rewarding its sourcing managers. Consequently the more suppliers the managers could persuade to achieve “next day” delivery, the higher would be their bonuses. The data for this bonus calculation was unreasonably not taken from the MRP system. But the
managers, realising this, set the lead-times in the MRP system to unity, despite not having actually agreed with the suppliers the new, reduced lead-time. In order to mark the effect of actual lead-times, the managers increased the safety stock. However, irrespective of the wrong doing becoming self evident via executive walk-about, our analysis suggests that this is an inadequate action. There is still a considerable danger that using such wrong information will result in an unstable (or at least a critically stable system therein), incurring yet more avoidable bullwhip on-costs.

Nor is such practice the only way of inducing chaos on the shop floor. We also have UK industrial experience of lead-time discrepancies occurring the other way around. Sometimes it is possible that a supplier, foreseeing a period of high workload in the near future, wants to deliver components early (but on a temporary basis). This may be due to requirements either to free up factory space or to reduce the burden or the transport system. An unstable system could result because the recipient company has built the shortened lead-time into the replenishment rule. But if there is close liaison between “players” in the value stream then this reduction would be flagged up as a transient expedient. It could then be treated as a known “interface” problem and be accounted for by appropriate action from the production scheduler. However, in contrast, the previously documented case of achieving bonus targets by manipulating data so as to benefit particular managers is an “internal” problem. As suggested earlier it can be detected by the Production Director going on a walk-about.

So on the basis of our investigations, what actions do we suggest in order to minimise risk of bullwhip induced reduced profits? A number of complementary actions are recommended as follows;

- Always monitor lead times automatically and update the replenishment rule accordingly. This is entirely within “our” remit.
- Strive to have speedy transparent high fidelity information flow speedily throughout the value stream. However it must be understood that considerable management effort and tenacity is needed to ensure that the right information actually flows between all “players”.
- Set the replenishment rule parameters very conservatively and well within our published stability contours. This will dampen down response to lead-time induced transients at the expense of a longer recovery period.
- Highlight lead-time changes so that when necessary a manual override is triggered. In other words the need for this action moves the product into the 5% “high activity” category from the 95% “automatic replenishing” group.

In short we need to establish both the likely operating scenario and the necessary IT support required to enable bullwhip risk to be minimised. This includes regularly exploiting simulation models to answer “what-if” questions by regularly forecasting product lead times when these might vary due to unavoidable and always changing complexities in shop floor scheduling and routing (Belk and Steels, 1998). Furthermore, adaptive replenishment rule parameter settings need to be related to current high fidelity information. Finally, to ensure that actual lead-times for all products are known to the production scheduler. Appropriately triggered intervention then determines if this particular product switches from automatic replenishment to “hands-on” mode.

6. Conclusion
Bullwhip is a costly phenomenon with many possible causes (Geary et al 2006). Furthermore it may be triggered anywhere within the supply chain. The project described herein is based on constraining one particular yet known common cause of bullwhip. This is the particular dynamic response induced by decision makers (usually people under great pressure) who consequently generate demand amplification via their “preferred” and often intuitive replenishment rule. Unfortunately such behaviour does not merely increase the risk of unnecessary on-costs arising for the incumbent echelon alone. As the Holmström (1997) study illustrates, there is a very good chance that, such a phenomenon will make matters much worse elsewhere in the supply chain. Other “players” will be wrong-footed and take what they consider to be appropriate compensatory action. Unfortunately the likelihood is that by now everything in the value stream is happening out-of-phase with each other. Thus, chaotic patterns of supply and demand become inevitable and can be disastrous for business.
Our theoretical investigations have already provided a neat DSS based theory for avoiding bullwhip in the matched lead-time case (Disney and Towill, 2006a). However if mismatch should occur in practice then bullwhip can unexpectedly appear. Under extreme conditions an unstable system can result. Hence our theory has now further evolved to cover this very demanding case. Consequently we are able to extend both theory and practice in terms of reducing bullwhip induced risk of profit losses via a number of suggested and tested actions. One of these strongly underpins the previous pragmatic approaches to the importance of transparent information flow proposed by Goldman et al, (1995) with reference to Carpenter Technology Inc. and by Christopher and Lee (2004). In other words it is just not enough to say that accurate lead-time estimates are first “a good thing to have”. Properly used they can greatly reduce on-costs arising from demand amplification.

Appendix A: System stability with mismatched lead times

The general form of a discrete transfer function can be written in z operator terms as

\[ F(z) = \frac{b_n z^n + b_{n-1} z^{n-1} + \ldots + b_1 z + b_0}{a_n z^n + a_{n-1} z^{n-1} + \ldots + a_1 z + a_0} \]  (A1)

The transfer function of our particular replenishment system that relates the order rate placed on our supply source to the demand rate received from our customer is given by

\[ F(z) = \frac{O(z)}{D(z)} = \frac{z^{1+T_p} (T_a(z-1)+T_p(z-1)+T_i(z-1)+z)}{(T_a(z-1)+z)(z^{T_p} + z^{T_p} ((1+T_i(z-1))z^{T_p} - 1))} \]  (A2)

The stability boundary of the system is determined by the denominator of this transfer function. For the purpose of determining system stability Disney (2007) highlights a criterion based upon matrices of the coefficients of the denominator that is very direct. In essence it is a trivial extension of Jury’s Inners well established approach, Jury (1974). It has the considerable virtue that it is extremely easy to automate with a computer. The method involves creating the following matrices,
The required system stability boundary is then given by the largest root of the matrix equation

$$a_n |X_{n+1} + Y_{n+1}| = 0.$$  \hspace{1cm} (A4)

Using such a procedure results $T_i > 0.5$ and $T_i > x$, where $x$ is calculated and listed in Table A1. This highlights the minimum value of $T_i$ required to ensure a stable response for different values of $T_p$ and $T_p$. Note that as befits our “hardware” engineering origins, we would traditionally opt for a well damped conservative design when setting parameter values (Towill, 1970). This means aiming to operate well away from the stability boundary. In this case a relatively large value of $T_i$ is required to achieve this goal, as already found experimentally in Table 2.

The range in stability requirements according to $T_p$ and $T_p$ shown in Table A1 is both surprisingly wide and indeed fluctuating. Note that the leading diagonal with $T_p = T_p$ requiring $T_i = 0.5$ confirms the validity of this particular boundary as previously established for DE-APIOBPCS by Disney and Towill (2003). What is surprising is how the boundary shifts around according to the relative values of $T_p$ and $T_p$. To visually make this point strongly, Figure A1 charts these results using $T_p$ as the baseline. The $T_p = T_p$ boundary is also highlighted.

The fluctuating behaviour of the limiting parameter values for stability is a manifestly a cause for considerable concern by the practising production controller. Of course in real-world business there is an extremely simple answer. But unfortunately it may not always be within individual company control. If the actual lead time $T_p$ is known for each product, and the estimate $T_p$ is continuously updated, then effective supply chain management and bullwhip control become much easier. The bullwhip benchmarking contours can then be used with much confidence. In practice there will be only a negligible risk of higher volatility than we have previously predicted (Disney and...

Towill, 2006a). So an investigation which initially studied bullwhip induced risk has resulted in yet stronger evidence supporting effective information exchange within the supply chain. Hence the narrative on this topic pragmatically advanced by Christopher and Lee (2004) is confirmed via transfer function analysis.

Figure A1. The minimum $T_i$ required for stability with mismatched lead-times
(Source: Authors)

<table>
<thead>
<tr>
<th>$T_i$</th>
<th>$T_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>1.414214</td>
</tr>
<tr>
<td>3</td>
<td>1.903212</td>
</tr>
<tr>
<td>4</td>
<td>2.073921</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>1.568512</td>
</tr>
<tr>
<td>7</td>
<td>1.821392</td>
</tr>
<tr>
<td>8</td>
<td>1.982066</td>
</tr>
<tr>
<td>9</td>
<td>2.064633</td>
</tr>
<tr>
<td>10</td>
<td>2.072477</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>1.91618</td>
</tr>
</tbody>
</table>

Table A1. Critical (minimal) values of $T_i$ required for stability for mismatched lead times
(Source: Authors)
Appendix B: Bullwhip and inventory variance expressions for the case of mismatched lead times

(a) General case

Using these previously determined matrices (A3) then for a stable system we may also estimate the bullwhip variance ratio from the equation

$$\text{Bullwhip} = \frac{\sigma_D^2}{\sigma_D^2} = \left[ X_{n+1} + Y_{n+1} \right]_b$$  \hspace{1cm} (B1)

where $\left[ X_{n+1} + Y_{n+1} \right]_b = \left[ X_{n+1} + Y_{n+1} \right]$ with the last row replaced by

$$[2b_0, 2\sum b_i b_{i+n-1}, \ldots, 2\sum b_i b_{i+1} + 2\sum b_i^2].$$

The corresponding inventory variance expression is given by

$$\text{InvVar} = \sigma_{NS}^2 = \frac{\left[ X_{n+1} + Y_{n+1} \right]_b}{a_n \left[ X_{n+1} + Y_{n+1} \right]}$$  \hspace{1cm} (B2)

where the coefficients now obviously relate to the transfer function of the inventory levels. This transfer function is

$$\frac{NS(z)}{D(z)} = \frac{z(\tau_i T_a(z-1) + z)^2 - T_i T_a(z-1)z^{\tau_i} + T_i(z-1)z^{\tau_i} + z^{\tau_i} - 1)}{(1 + z)(T_a(z-1) + z)^2 - z^{\tau_i} + T_i(z-1)z^{\tau_i} + T_i z^{\tau_i} - T_i - 1).}$$  \hspace{1cm} (B3)

(b) Specific cases when $T_p \neq T_p$

This procedure further requires us to specify a value of $T_p$ and $T_p$, but this can also be computed even for very high order systems. For the particular case when $T_p=1$ and $T_p=3$ then the bullwhip expression becomes

$$\text{Bullwhip} = \frac{\sigma_D^2}{\sigma_D^2} = \left[ 2T_a^2 (T_i - 1)(1 + T_i^2) + T_a^2 (T_i - 1)(23 + 6T_i)(1 + T_i^2) + \frac{T_i T_a(49 + T_i(52 + T_i(19 + 2T_i)))+ T_a^2 (T_i (122 + 57T_i + 6T_i^2) - 104) - 49}{T_a^2 (T_i (20 + T_i(14 + T_i(62 + T_i(31 + 2T_i)))) - 97)} \right]$$  \hspace{1cm} (B4)

which is arrived at by working through the ensuing tedious, but straightforward algebra. The corresponding inventory variance expression when $T_p=1$ and $T_p=3$ is
When \( T_p = 3 \) and \( T_p = 1 \) the following expressions for bullwhip and inventory variance are obtained.

\[
\text{Bullwhip} = \frac{\sigma_{\text{bullwhip}}^2}{\sigma_D^2} = \frac{(T_i + T_p) (T_i + T_p + T_p + 2T_p) + (2T_i + T_p + 2T_p + 3T_p) + (7 + T_p + 2T_p + 27 + T_p + 41 + T_p + 29 + 6T_p))I^4 + (1 + T_p)^4 T_i}{(1 + 2T_p)(T_i - 1)\{1 + T_i\}(2T_i - 1)\{1 + T_i\}(2T_i + 3T_i + 2T_p + T_p)^3 + (1 + T_p)^4 T_i}.
\]

(B5)

\[
\text{InvVar} = \frac{\sigma_{\text{InvVar}}^2}{\sigma_D^2} = \frac{(T_i + T_p) (T_i + T_p + T_p + 2T_p) + (2T_i + T_p + 2T_p + 3T_p) + (7 + T_p + 2T_p + 27 + T_p + 41 + T_p + 29 + 6T_p))I^4 + (1 + T_p)^4 T_i}{(1 + 2T_p)(T_i - 1)\{1 + T_i\}(2T_i - 1)\{1 + T_i\}(2T_i + 3T_i + 2T_p + T_p)^3 + (1 + T_p)^4 T_i}.
\]

(B6)

(b) The special case of equal lead times \( T_p = T_p \)

In the previously evaluated case of \( T_p = T_p \) then the closed form bullwhip estimate becomes

\[
\text{Bullwhip} = \frac{\sigma_{\text{bullwhip}}^2}{\sigma_D^2} = \frac{2 T_i^2 + 3T_i + 2T_p + 2(T_i + T_p)^2 + T_p (1 + 6T_i + 4T_p)}{(1 + 2T_i)(T_i + T_p)(-1 + 2T_i)}.
\]

(B8)

Equation (B8) is especially interesting as it holds for all lead-times (when \( T_p = T_p \)).

As does the inventory variance expression. It is

\[
\text{InvVar} = \frac{\sigma_{\text{InvVar}}^2}{\sigma_D^2} = 1 + T_p + \frac{2 T_i^2 (T_i - 1)^2 + T_i (1 + T_p)^2 + T_p (1 + T_p) (1 + (2T_i - 1)I^2)}{(1 + 2T_i)(T_i + T_p)(2T_i - 1)}.
\]

(B9)

References


