Removing bullwhip from Lexmark’s toner operations

Stephen M. Disney (*disneysm@cardiff.ac.uk*)
Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK.

Lance Hoshiko (*lhoshiko@lexmark.com*)
Lexmark International, 6555 Monarch Road, Longmont, CO 80503, USA.

Lorin Polley (*lpolley@lexmark.com*)
Lexmark International, 6555 Monarch Road, Longmont, CO 80503, USA.

Courtney Weigel (*weigelc@lexmark.com*)
Lexmark International, 6555 Monarch Road, Longmont, CO 80503, USA.

Abstract
Lexmark’s existing planning system, that controls the production of printer toner, created a bullwhip effect of 5.4:1. Inventory variability, measured as a ratio to the demand variance, was 27:3. We describe a project where, since 2011, we have been able to reduce bullwhip to 0.94:1 and inventory variability to 2.15:1.

Keywords: Bullwhip, Inventory, Production Planning

Introduction
In early January 2010, Lorin Polley, Lexmark’s Boulder Site General Manager, was reading Disney and Lambrecht (2008)… “the bullwhip effect can be quite exasperating for companies; they invest in extra capacity, extra inventory, work over-time one week and stand idle the next, while at the retail store the shelves of popular products are empty and the shelves with products that are not selling are full”. It occurred to him - this is exactly the problem we have here in Lexmark. Lorin telephoned Stephen Disney, one of the authors of the article he was reading. After three years of fortnightly teleconferences, a nine week secondment to the Boulder Lexmark facility, and a three day executive training program, we have made progress at reducing the bullwhip problems in Lexmark’s toner operations. However there are still challenges to be tackled and Lexmark’s battle with bullwhip is not yet over. But this is the story so far…

Understanding the Lexmark Supply Chain: Value Stream Mapping
Our initial discussions focused on understanding and documenting Lexmark’s toner supply chain. In order to do this we exploited the Value Stream Mapping of Rother and Shook (1998) to chart both the material flow and the information in Lexmark’s printer

toner operations. Overlaid on the VSM we collected time series data of demands, inventory and orders, which allowed us to see where in the supply chain bullwhip problems were being generated. We have not shown the diagram in order to save space and to preserve confidential information. However, for an example real product, we have provided has some example time series in Figure 1 with a bullwhip measure of $Bullwhip = \left(\sigma^2_{Orders} / \sigma^2_{Demand}\right) = 6.7$ and $NSAmp = \left(\sigma^2_{Net\ Stock} / \sigma^2_{Demand}\right) = 16$.

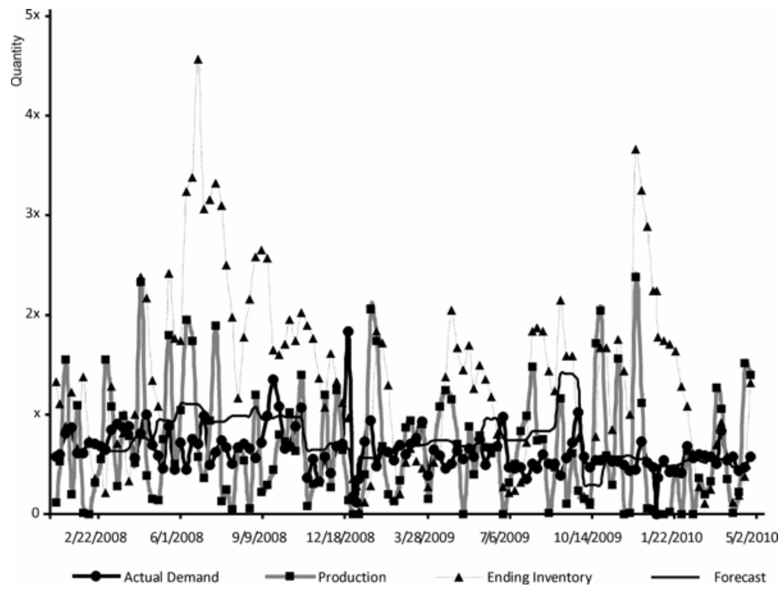

Figure 1. Example time series of demand, production, inventory and forecasts before our project

Identifying the root causes of the bullwhip problem in Lexmark: Ishikawa diagram

In conjunction with the VSM map we also held a brainstorming session with Lexmark's Operational Excellence Team and generated an Ishikawa diagram of the root causes of the bullwhip effect in Lexmark, see Figure 2.

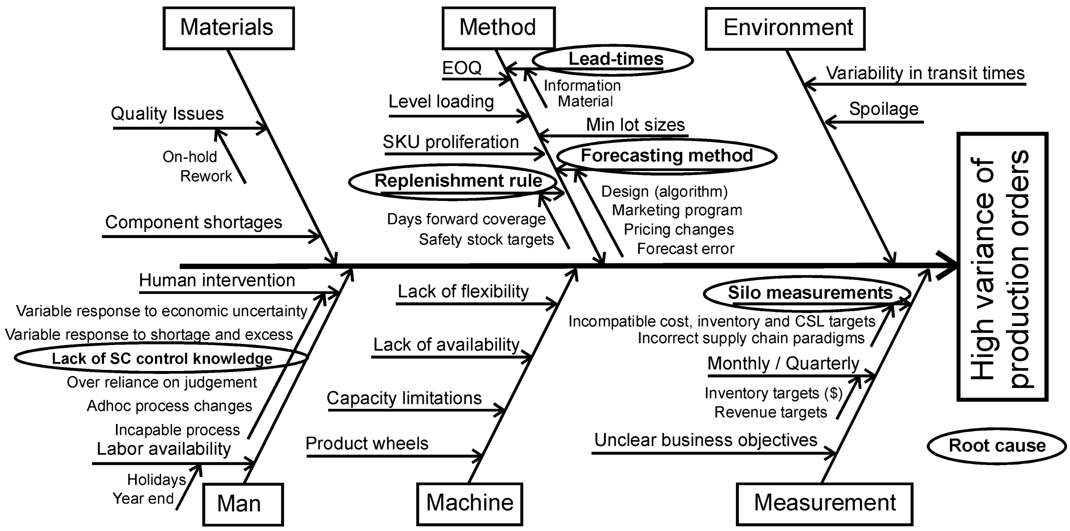

Figure 2. Ishikawa diagram of the causes of the bullwhip effect within Lexmark

It is worth discussing some of these root causes in more detail. We will focus on the forecasting and the replenishment rule issues here.

Forecasting at Lexmark

Lexmark has their own forecasting team, generating forecasts for individual cartridges in each of their geographical markets. These forecasts are based on the current (and predicted) estimate of the number of printers in the market place and an average "page count". This forecast is then subjected to a judgemental adjustment based on market knowledge. When we aggregated these forecasts into a forecast of the demand for individual toners, we found that there was both a high forecast error and, at times, severe bias. In short, it was our opinion that the current Lexmark forecast was not suitable for use as a forecast for production and distribution planning.

Existing replenishment rules in Lexmark

When we started our project in 2010, Lexmark were using a "Day's Forward Coverage" replenishment policy. In essence, this policy would take the current forecast for a product and set the safety stock levels to a multiple of the forecast. To that it would create a replenishment order by comparing the inventory position to this dynamic safety stock and adding the difference to a forecast of demand over the lead-time and review period. This was an extremely volatile system and results in the extreme levels of bullwhip that we can see in Figure 1.

Lean production and the introduction of SAP in Lexmark

During 2010 Lexmark introduced SAP, going live on 2nd August 2010. With this IT change, the "Days Forward Coverage" replenishment policy was replaced with an "Order-Up-To" (OUT) policy. See (1) for a common formulation of the OUT policy,

$$o_t = \underbrace{\hat{d}_{t+T_p+1,t}}_{\text{Forecast}} + \underbrace{\overbrace{tns}^{\text{Safety stock}} - ns_t}_{\text{Inventory feedback}} + \underbrace{\overbrace{\sum_{i=1}^{T_p} \hat{d}_{t+i,t}}^{\text{DWIP}} - \overbrace{\sum_{j=1}^{T_p} o_{t-j}}^{\text{WIP}}}_{\text{WIP feedback}} \tag{1}$$

In (1), o_t is the orders placed at time *t*. $\hat{d}_{t+T_p+1,t}$ is the forecast of demand, made at time *t*, for the demand in the period $t+T_p+1$ where T_p is the lead-time (the "+1" is the review period required to ensure the correct sequence of events). *tns* is the target net stock, a time invariant constant set to ensure an appropriate level of inventory availability. In reality the *tns* was (re-)set every quarter to reflect the latest demand profile. ns_t is the current net stock level at time *t*. The work-in-progress (WIP) is the sum of the open orders that have been placed, but not yet received. The desired WIP is the sum of the forecasted demand over the lead-time.

Another significant change happened in Lexmark's cartridge assembly plants in 2010. With the introduction of Heijunka Boards and a large scale Lean Production Initiative, the cadence of the pacemaker controlling the level of the cartridge assembly was set to customer demand. This change allowed a much smoother demand signal to be placed onto the toner production plant. This production smoothing regime was completely in line with the advice of Ohno (1988) who stated that "The slower but consistent tortoise causes less waste and is much more desirable than the speedy hare that races ahead and then stops occasionally to doze. The Toyota Production System can be realized only when all the workers become tortoises".

The Toner Production Planning Pilot

In order to allow Ohno's vision to be followed in the Toner factory supplying the cartridge assembly plants we developed a spreadsheet application (TonerPPP.xls) during May-July 2011. This spreadsheet took data from SAP and made recommendations to the Toner Production Planner on how much toner to produce and ship to the assembly plants. We will now explain the major design choices made and how the system works.

The demand profile over the lifetime of a toner was found to broadly follow a 10-15 year pattern. When the toner was "new" and was being used in printers currently being sold in the marketplace, demand for toner was found to generally increase linearly. Over time though, a new toner would be developed and new printers would use the new toner. However, the existing printers in the market would still be consuming the original toner. In this period, demand for a toner was reasonably level. Eventually however, the printers in the market would be replaced. During this period, (which may last 5-10 years) the demand for toner was following a slow exponential decay. There was also a random (noise) component and a quarterly "hockey stick" effect overlaid onto this broad pattern.

These facts, together with an understanding of the company lead-times and known desirable characteristics of forecasts from a supply chain perspective led us to use the Damped Trend forecasting mechanism within our TonerPPP.xls. The Damped Trend (Gardner and McKenzie, 1985) forecast is given by

$$\left.\begin{aligned} \hat{a}_t &= (1-\alpha)\left(\hat{a}_{t-1}+\phi\hat{b}_{t-1}\right)+\alpha d_t \\ \hat{b}_t &= (1-\beta)\phi\hat{b}_{t-1}+\beta\left(\hat{a}_t-\hat{a}_{t-1}\right) \\ \hat{d}_{t,t+k} &= \hat{a}_t+\hat{b}_t\sum_{i=1}^{k}\phi^i \end{aligned}\right\} \quad (2)$$

Here d_t is the demand at time *t*. $\hat{a}_t$ is the current estimate of the level of the demand, exponentially smoothed by the constant level α. $\hat{b}_t$ is the current estimate of the trend, exponentially smoothed by the constant β. $\hat{b}_0$ is the initial value of the trend, assumed to be zero, $\hat{b}_0 = 0$. ϕ is a damping parameter that can be tuned to account for the persistence of the trend. k is the number of periods ahead that the forecast is required to predict, thus $\hat{d}_{t,t+k}$ is the forecast, made at time t, of the demand in the period $t+k$. (2) suggests that the estimate demand consists of time dependant "constant" and "trend" components.

Several well-known forecasting approaches are encapsulated within the damped trend forecasting model (Gardner and McKenzie, 2011). These include Holt's method where there is no damping of the trend component, $\phi = 1$, Simple Exponential Smoothing (SES) when $\beta = 0$ and Naïve forecasting when $\alpha = 1$ and $\beta = 0$. In our experience we found Holt's Method setting (of $\phi = 1$) resulted in the most useful forecasts for toners. It is also available native in SAP, aligning the TonerPPP.xls design to a potentially "easy to implement" situation in SAP.

The next choice that we made was to alter the OUT replenishment rule. The fact that the OUT policy creates bullwhip is no surprise to a control engineer - there are unit feedback gains in the feedback loops, and this will create a under-damped system. Indeed, Dejonckheere et al (2003) has shown that the OUT policy with moving average and exponential smoothing forecasts will always generated bullwhip, for all demand patterns, for all replenishment lead-times. Li and Disney (2012) has extended this result by showing that Holt's Method will also always generate bullwhip, for all lead-times and all possible demand patterns.

An experienced control engineer would modify the OUT policy by introducing proportional controllers into the inventory and WIP feedback loops. This is done by

$$o_t = \hat{d}_{t+T_p+1,t} + \tfrac{1}{T_i}\left(tns - ns_t\right) + \tfrac{1}{T_w}\left(\sum\nolimits_{i=1}^{T_p} \hat{d}_{t+i,t} - \sum\nolimits_{j=1}^{T_p} o_{t-j}\right), \qquad (3)$$

where there are two independent feedback controllers in the net stock and WIP feedback loops. We call (3) the Proportional OUT policy. The proportional feedback controllers, T_i and T_w can be tuned to give the system response certain dynamic characteristics, Disney and Towill (2003). Control engineers have built up a vast range of mathematical tools to design such systems as this type of problem occurs in many physical systems. For example, we see similar control problems when regulating the speed of stream engines, the temperature of furnaces, in automotive cruise controls, ship steering systems, aircraft auto-pilots to name few. Care has to be taken when setting T_i and T_w as in extreme cases, inappropriate settings can result in unstable and chaotic systems, Wang et al. (2012). However, when $T_w = T_i$, the system has a number of desirable characteristics. For example, it is stable for all $T_i \in \Re > \tfrac{1}{2}$, and the orders are robust to both random production yields and stochastic lead-times when the average lead-time is correctly identified, Disney and Towill (2001).

The $T_w = T_i$ setting also results in a very simple set of *Bullwhip* and *NSAmp* expressions when demand is i.i.d. and optimal forecasts are used, see (4) and (5).

$$Bullwhip = \frac{\sigma^2_{Orders}}{\sigma^2_{Demand}} = \frac{1}{2T_i - 1} \qquad (4)$$

$$NSAmp = \frac{\sigma^2_{NetStock}}{\sigma^2_{Demand}} = 1 + T_p + \frac{\left(T_i - 1\right)^2}{2T_i - 1} \qquad (5)$$

We have plotted (4) and (5) in Figure 3 *Bullwhip* increases from zero when $T_i \to \infty$, to unity when $T_i = 1$ and to infinity when $T_i = \tfrac{1}{2}$. The *NSAmp* is infinite when $T_i = \left\{\infty, \tfrac{1}{2}\right\}$ and has a minimum of $1 + T_p$ when $T_i = 1$. If bullwhip (capacity) and inventory costs were equal and proportional to the variance of orders and the net stock, then the minimum costs would be incurred by setting $T_w = T_i = 1.618034$, the golden ratio, Disney et al (2004). Sadly however, inventory and capacity costs turn out to be a function of the standard deviation of the orders and inventory levels and the elegant answer, the golden T_i, does not hold. The optimal $T_i > 1$, but it is not necessarily the case that the optimal $T_i = 1.618034$.

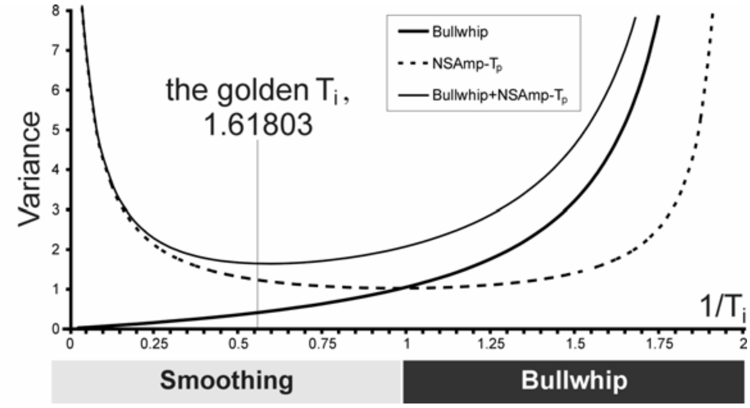

Figure 3. Bullwhip and NSAmp in the Proportional OUT policy with i.i.d. demand

We incorporated (3) into the TonerPPP.xls. We used one worksheet for each type of toner. We allowed the replenishment policy to dynamically update forecasting parameters, feedback controller and safety stock targets each week for each toner at all locations. A summary input / output diagram, as well as a systematic of a supply chain is shown in Figure 4.

The strategic planning activity

To set-up (and re-tune periodically, perhaps every quarter) the safety stock target, forecasting and POUT policy parameters we created a "POUT policy designer" worksheet within the TonerPPP.xls. When an individual toner is selected, a macro gathers demand, forecast inventory, distribution and production information from around the spreadsheet. This information is then used to create a "what if" simulation of what could be achieved if particular parameter settings where chosen. A range of supply chain and financial performance measures are calculated (tackling one of the root causes of the bullwhip effect from Figure 2) to aid this decision. When a desirable set of parameters are found in the policy designer, a macro re-distributes these parameters back into the "live" worksheets. A screenshot of the OUT policy designer is shown below in Figure 5.

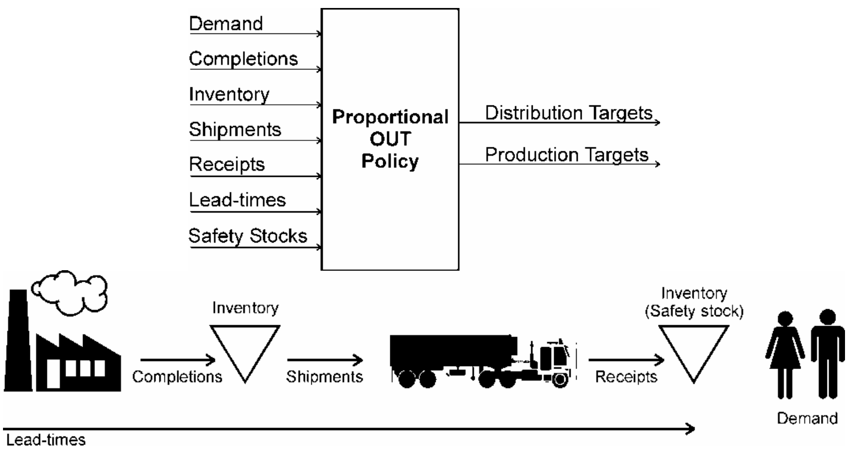

Figure 4. The information used by the Proportional OUT policy

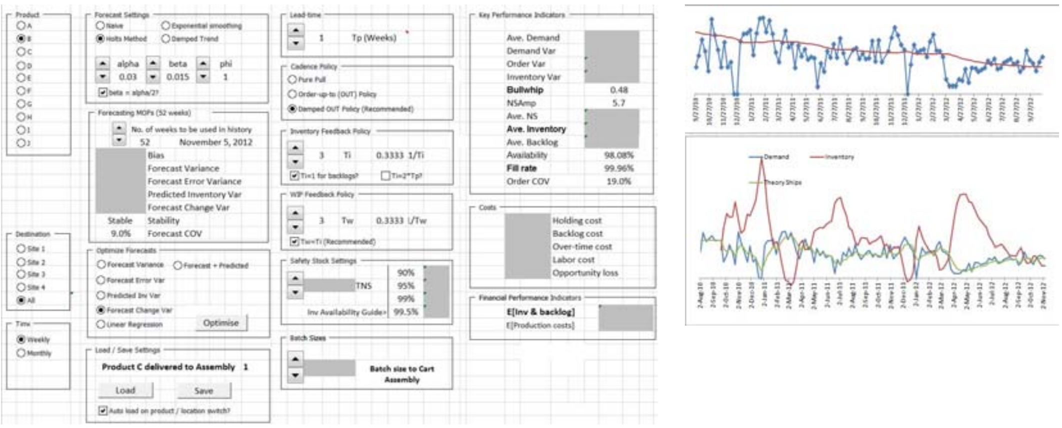

Figure 5. Screenshot of the OUT policy designer

The weekly production planning activity

On a weekly basis the TonerPPP.xls worksheet is used to guide the production and distribution planning activity. Each Sunday night, the Production Planner gathers information from SAP concerning: the usage of Toner at the cartridge assembly sites; the local inventory (raw materials) of toner at the cartridge sites; the quantity that was received during the previous week; and the quantity that was dispatched to those locations from the toner factory. The Production Planner also gathers information about the inventory of toner held at the toner factory (as finished goods) and what was produced each week. This information can all be gathered by SAP electronically and a report generated (as a *.csv file) automatically if desired. This *.csv file can then be imported into the TonerPPP.xls spreadsheet automatically with the aid of a macro. However, we have found that it is better to enter the data into the spreadsheet manually. As there are a limited number of different toners are planned in this mode, the manual input means that some “expert eyes” are passed over the data enabling validation. For example, if two types of toner are put on a truck (or in a container) to an assembly plant, and one of those toner types is recorded as having arrived, and the other has not, the production planner knows that there is a problem with the data in SAP. This type of tacit knowledge is essential, even with such a sophisticated ERP package as SAP.

After the data are input into the spreadsheet a macro distributes the data to individual toner worksheets within the spreadsheet file. Using the current setting for safety stock targets, forecasting parameters, lead-times and feedback controllers those individual worksheets determine how many products to distribute to each location and then, after netting any remaining inventory at the factory from the previous week, what quantity of toner to produce in the factory. As toner is packaged in bulk sacks, distribution and production targets are rounded up to fill a sack.

These distribution and production targets are then reviewed by the Production Planner, see Figure 6. The Production Planner can, at this stage, impose a Minimum Order Quantity on particular toners in order to avoid lost capacity and excessive changeover costs. The Production Planner may also exploit tacit knowledge about the supply chain here. For example, recently a wildfire prevented transportation along a standard route, thus dispatches were delayed. If a machine were down for maintenance, perhaps particular toners could not be produced. Planning around National holidays in the

Cartridge Assembly 1					
Product	Disney Shipments	Confirm	Alternative Shipments	Current Inventory	Last Weeks Consumption
A	0	☑ Approve?		6,484	2,143
B	32,200	☑ Approve?		38,142	24,104
C	22,425	☑ Approve?		9,961	13,993

Cartridge Assembly 2					
Product	Disney Shipments	Confirm	Alternative Shipments	Current Inventory	Last Weeks Consumption
A	0	☑ Approve?		4,578	726
B	6,860	☐ Approve?	5,880	23,515	5,357
C	4,900	☐ Approve?	0	23,931	4,947

Toner Manufacturing						
Product	Disney Production Orders	Confirm	Alternative Production	Ending Inventory	MOQ	
A	0	☑ Approve?		0	☐	15,306
B	16,378	☐ Approve?	25,500	9,122	☐	44,887
C	16,421	☐ Approve?	20,000	3,579		

Figure 6. Confirming the distribution and production targets in the TonerPPP.xls

USA and the countries of the assembly plants can also be accomplished.

Performance of the TonerPPP.xls

The TonerPPP.xls was developed, implemented and has been running since May-July 2011. The spreadsheet system also serves as a database of demand, inventory, delivery and production information and planning decisions. Thus we are able to interrogate this database to determine the success of the project. See Figure 7 for a typical visualization of bullwhip performance for a single toner. On the x-axis we have plotted time in the weekly planning cycles. The y-axis captures a rolling 52 week variance ratio. We have plotted three variance ratios. The first is the ratio of the variance of the TonerPPP.xls recommended production order to the variance of demand. We noted that this ratio never quite achieves the variance ratio that we would expect from our theoretical models. This is because the supply chain never quite “sticks to the recommended plan”. Any deviations from the system recommendation, or indeed the agreed (by the Production Planner) plan and what was actually dispatched to assembly sites and manufactured by the factory, will eventually turn up in WIP and inventory states. The system does attempt to recover from these deviations in a damped, considered manner, but this is a source of additional noise in the real world that is not present in the “academic” models.

The second ratio tracks the variance of the agreed production order to the demand variance. One can see that the Production Planner allows some additional variation creating more bullwhip into the system, but presumably this is driven by tactical knowledge. For example, dealing with disruptions and wider supply chain issues is an important and vital task of the Production Planner, and one that can’t readily be implemented into a spreadsheet.

The third ratio tracks the variance of the actual weekly production completions to the demand variance. We can see that the TonerPPP.xls is producing a production order with less variance of demand, and an agreed production plan is produced with less variance than demand in the second half of 2012 (when the old production system falls out of the data window in the 52 week rolling horizon). However, the factory is unable to consistently perform to the agreed plan. This creates a residue amount of bullwhip, that we are having trouble eliminating and is an important difference to the bullwhip project in Tesco that was previously reported in Potter and Disney (2010). In Tesco’s retail supply chain, very heavy penalties for late or incomplete deliveries were in place, so this issue with “sticking to the agreed plan” did not become a significant issue. Common reasons difficulties of sticking to the agreed plan include; overreacting to downstream communications, equipment availability issues, or volatility during new product ramp up. Also, the quarterly updating of the system parameters introduces a further source of the

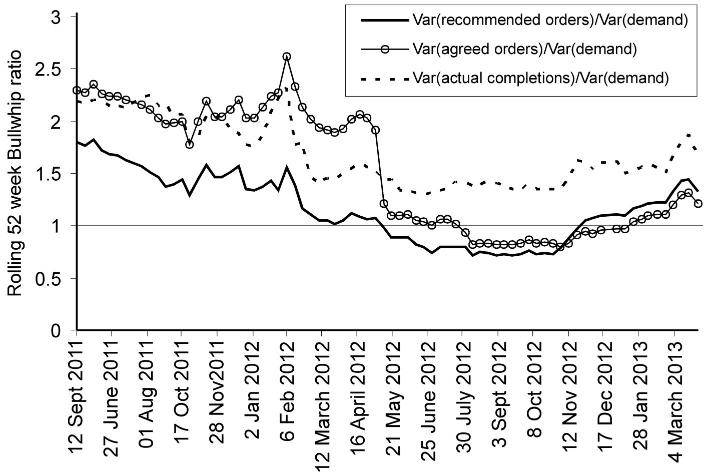

Figure 7. Rolling 52 week bullwhip ratios for an individual toner

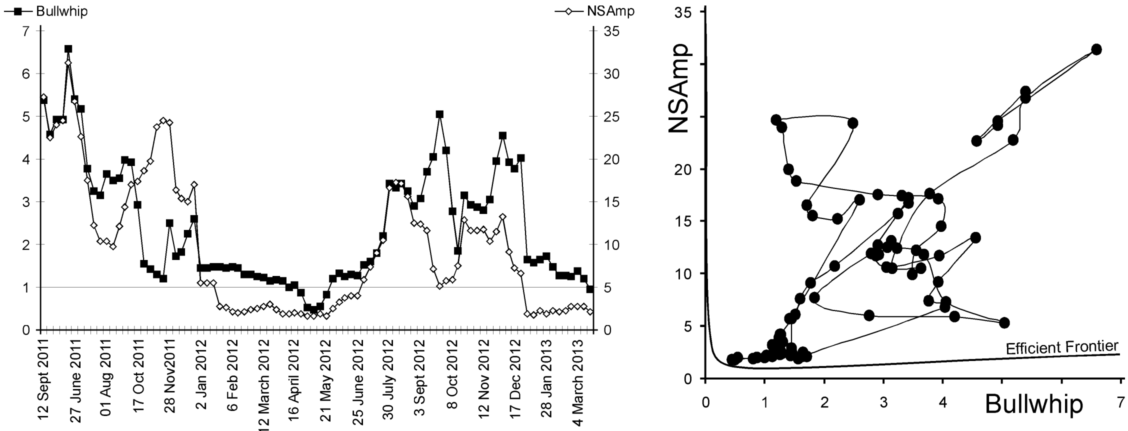

Figure 8. Bullwhip and inventory performance at Lexmark (based on 19 week rolling ratios)

bullwhip effect.

Another way to investigate the performance of the TonerPPP.xls is to plot the *NSAmp* and *Bullwhip* (actual factory completions) performance on an Efficient Frontier graph, see Figure 8. Here the efficient frontier was generated from (4) and (5) (with $T_p = 0$). We can see from Figure 8 that, across the whole toner factory, we have been able to reduce *Bullwhip* (here measured on a 19 week rolling basis) from 5.4 to 0.94 and *NSAmp* from to 27.3 to 2.15. During the latter half of 2012, factory bullwhip resurfaced due to a number of tactile issues. However, recently the dynamic performance has been brought back into line.

The future of Bullwhip at Lexmark

Our current plans for our Bullwhip project are to implement the Proportional OUT policy into SAP as a user exit function. The formulation of the POUT given in (6) looks promising as it only requires demand, forecast and order information and thus may be easier to implement in SAP. This may also allow us to design a replenishment policy that is impervious to the actual execution of the production plan – giving another benchmark to the Production Planner when making replenishment decisions.

$$o_t = \hat{d}_{t+T_p+1,t} + \frac{1}{T_i}\left(\left(\sum_{i=1}^{T_p+1}\hat{d}_{t+i,t}\right) - \left(\sum_{i=1}^{T_p+1}\hat{d}_{t+i-1,t-1}\right) + d_t - \hat{d}_{t+T_p+1,t}\right) + \left(\tfrac{T_i-1}{T_i}\right)\left(o_{t-1} - \hat{d}_{t+T_p,t-1}\right) \qquad (6)$$

A management and control system that monitors adherence to the production plan is also required. The production plan adherence is affected by equipment, manpower and material availability. The source of the increased bullwhip may also be due to changing safety stock targets. We are currently investigating options available to cope with this internally generated variability.

Acknowledgements

We would like to thank the many people in Lexmark who have contributed to our bullwhip project. There were many people involved, but Elizabeth Bartilson, Mick Dulany, Rebecca Silveston-Keith, DuWayne Zuehlsdorff and Scott Mosier deserve special mention.

Stephen would also like to especially thank Lorin Polley, Lance Hoshiko, Courtney Weigel, Kathy Rayman and Trish Fitch for making the Disney family's adventure to Colorado so enjoyable – it was a very special time for us, thank you.

References

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

Disney, S. M., D. R. Towill. 2001. A linear PIC system that can operate in a non-linear world. *Proceedings of the IFAC MIM 2001 Conference in Manufacturing, Modeling, Management and Control*, 2-4 August, Prague, Czech Republic.

Disney, S. M., D. R. Towill. 2003. On the bullwhip and inventory variance produced by an ordering policy. *OMEGA: The International Journal of Management Science* **31**(3):157-167.

Disney, S. M., M. R. Lambrecht. 2008. On replenishment rules, forecasting and the bullwhip effect in supply chains. *Foundations and Trends in Technology, Information and Operations Management* **2**(1): 1-80.

Disney, S.M., D. R. Towill, W. Van de Velde. 2004. Variance amplification and the golden ratio in production and inventory control. *International Journal of Production Economics* **90**(3): 295-309.

Gardner, E.S. Jr, E. McKenzie. 1985. Forecasting trends in time series. *Management Science*, **31**(10): 1237-46.

Gardner, E.S. Jr, E. McKenzie. 2011. Why the damped trend works. *Journal of the Operational Research Society*, **62**(6): 1177-80.

Li, Q., S. M. Disney. 2012. Damped Trend forecasting and the Order-Up-To replenishment policy. *Pre-prints of the 17th International Working Seminar of Production Economics*, Innsbruck, Austria, February 20th -24th, **3**: 205-216.

Ohno, T. 1988. *Toyota Production System: Beyond large scale production.* Productivity Press, Portland, Oregon, USA.

Potter, A., S. M. Disney. 2010. Removing bullwhip from the Tesco supply chain. *Production and Operations Management Society Annual Conference,* May 7th - 10th, Vancouver, Canada. Paper No. 015-0397, 19 pages.

Rother, R., J. Shook. (1998). *Learning to See: Value Stream Mapping to Add Value and Eliminate Muda*. Lean Enterprise Institute, Brookline, MA, USA.

Wang, X., S. M. Disney, J. Wang. 2012. Stability analysis of constrained inventory systems with transportation delay. *European Journal of Operational Research* **223**(1): 86-95.