Understanding your supply chain: dynamic value stream mapping for business improvement

Xia Meng
Cardiff Business School, Cardiff University, UK

Stephen M. Disney
Cardiff Business School, Cardiff University, UK

Laura Purvis (purvisl@cardiff.ac.uk)
Cardiff Business School, Cardiff University, UK

Abstract

The "Bullwhip effect" is a well-studied phenomenon in SCM. However, despite the large amount of work in the field, mainly simulation based, the number of empirical studies exploring the causes of bullwhip behaviour and their impact is limited. Our paper presents a study of a dairy manufacturer with a distinctive set of characteristics: short shelf life, weekly heartbeat-shaped demand, a multi-product wheel with limited production capacity, short lead-time requirements. By using dynamic Value Stream Mapping, we demonstrated the extent to which poor SC dynamic behaviour can create and amplify the bullwhip effect. We then explore the causes of this behaviour and identify opportunities for improvement.

Keywords: Value stream mapping, System dynamics, Bullwhip effect

Introduction

The "Bullwhip effect" is recognised as a typical phenomenon in the field of operations and SCM. It occurs when production order variance amplifies with increases in demand variance (Lee et al., 1997). The problems associated with the Bullwhip effect can be economically and operationally costly, because they lead to excessive stock holding, order backlogs, late deliveries, under/over resource utilization. Theoretically, bullwhip reflects the existence of various inefficiencies in operations and SC processes, as well as Muda (waste), Mura (overburden), Muri (unevenness), which are the three key lean concepts established by Ohno (1998).

To date, the number of primary empirical studies that provide evidence on how bullwhip reveals within certain SC settings and how it can be mitigated against through optimal SC behaviour, decisions and strategies is very limited. This is a gap that our study aims to fill.

Literature Review

The topic of Supply Chain (SC) Dynamics has attracted academics' and practitioners' attention for over five decades, ever since Forrester (1961) discovered that demand fluctuations amplify as orders flow upstream. Also known as the 'bullwhip effect', the phenomenon creates excessive stock holding, backlogs, late deliveries, under/over resource utilization (Lee et al., 1997). Follow up work has also highlighted that attempts to improve poor SC dynamic behaviour can actually exacerbate the problem. Counter-intuitive behaviour often occurs because the causes of the behaviour are obscured from the decision-makers in the SC (Naim et al., 2004).

The large majority of studies on the Bullwhip effect has been conducted within an experimental setting in order to, for example, understand its root causes (Ghali, 2003) and examine the extent to which participants' decisions cause bullwhip related problems (Croson et al., 2013). The "Beer Game" has been the most employed approach, ever since established by Sterman (1989). In many Beer Game experiments (e.g. Cao et al., 2016), participants were observed to be irrational, which led to poorly informed decisions on production orders and replenishments. Participants' behaviours also varied when the objectives changed and the cost structures were different.

Several empirical studies have also observed the existence of the bullwhip effect in practice (e.g. Isaksson and Seifert, 2016). However, where empirical data is used, the assumptions that are built into the different modes employed vary – for example, different criteria are set to define a rational decision; or publicly-available, rather than purposely collected data is used to examine bullwhip at industry-level (Cachon et al., 2007). As a result, the practical underpinning of the theory associated with the bullwhip effect may not be as solid as it could be. Wang and Disney (2016) further emphasise that research could benefit from empirical, experimental and analytical approaches to identify the real cost structure as well the typical and optimal dynamic responses. Specifically, case studies and empirical analysis based on firm-level rather than industry-level, data have the ability to offer more insights on the incentives of demand smoothing and amplification. Our study aims to fill this gap.

Study Design

A single, longitudinal case study was conducted with one of the UK's largest dairy manufacturers. The uniqueness of their operations and SC systems stems from the fact that, first, their products are perishable, with a very short shelf-life, while also being very cost sensitive. As a result, identifying an optimal batch size that achieves the best trade-off between production costs and reduced levels of obsolescence is critical. Large batch size can indeed minimise fixed production wastage and reduce machine changeover time, but can potentially result in obsolescent products, where products are 'too old' to be accepted by retailers. Also, if demand was over-estimated, excessive inventory will result, with higher risk of obsolescence. However, small batch sizes are not cost-efficient as the percentage of fixed production wastage can be too high when too many change-overs occur.

Second, retailers place orders every morning and expect products to be despatched by late afternoon, with products arriving at their warehouses by late night. The expected service level is in excess of 99%, and when this target is missed, penalties can be incurred. With production times longer than the expected lead-times, the case company has to plan production based on forecasted rather than real demand. Interestingly, if it happens that the company's opening stock is not enough to cover the expected demand, they may decide on a "high-priority" production run – which they name *"Needed for tonight"* (NFT). However, this is risky action, as production and despatch that happen on the same day has a very tight schedule due to limited spare capacity, so there is a chance they will be late for the despatch deadline. Despite this risk, they may still attempt to create NFT if they want to maintain high service levels.

Third, as mentioned above, production is always planned according to demand forecasts. Production planning requires high level of forecast accuracy and/or optimal safety stock in order to reduce the number of possible NFTs. The manufacturer produces a large variety of dairy products with drastically different demand patterns, thus forecasting can be challenging for some SKUs.

Fourth, the limited production capacity is seen as a main constraint when it comes to production smoothing. Other constraints include: minimum and maximum levels of dairy base-making (sub-component of finished product) can restrict production optimisation. Machine speed can also limit the maximum production capacity for finished products. The production of some ranges of products is guided by a product wheel, for example some need to be produced on the same day and on the same machine – this requires optimisation of raw material, machine and labour capacity. Limited by all these constraints, the company aims to dispatch fresher products and reduce its costs while increasing its sales.

Data collection centred on the company's forecasting, production planning and inventory control activities along 9 value streams (products) in order to explore the causes of dynamic behaviour and their impact at various nodes within the system, over a period of 12 months. The choice of 9 products covers combinations of various product characteristics, such as sales volumes, promotion activity frequency, and whether the case company directly employs forecasts provided by their customers (retailers) or generates its own. A mixed methods approach was employed, combining quantitative statistical analysis and modelling with qualitative interviews and focus groups involving top and middle-managers across various functions of the organisation.

First, the study employed value stream mapping (VSM) to explore the material and information flows within the business – from the initial customer, through production planning, to delivery. We also mapped the generation of supplier orders, based on the company's forecasted requirements, through supplier deliveries to raw material inventories. The VSM is particularly useful as it visualises how material and information flows interact (Rother and Shook, 1999). For each of the levels and flows in the VSM, time series data was collected and analysed (Figure 1), where a number of measures were used: NSAmp ratio (Disney and Lambrecht, 2008), bullwhip ratio (Lee et al., 1997), fill-rate, forecast accuracy (Wackerly et al., 2008) and forecast nervousness (Li and Disney, 2016). The NSAmp ratio is applied at each of the inventory locations in the value stream, and is a useful measure of the company's ability to manage its inventory. The variance of the production orders divided by the variance of the demand is known as the Bullwhip ratio (Li et al., 2014). To measure forecast accuracy, we choose the mean squared error over the lead-time and review process, as this is directly related to the inventory variance, finished goods NSAmp, and safety stock requirements. Forecast nervousness measures the accuracy of future forecasts streams (Li and Disney, 2016). The measure assigns more weight to near-term forecasts errors, less so for forecast errors further into the future. These future forecasts are particularly important for raw material procurement when companies plan over a longer lead-time, particularly when the supplier has a long lead-time.

Findings

We present our study's findings following Wang and Disney (2016)'s approach, who include as main elements for bullwhip modelling: demand forecasting, time-delays, ordering policies and information sharing practices.

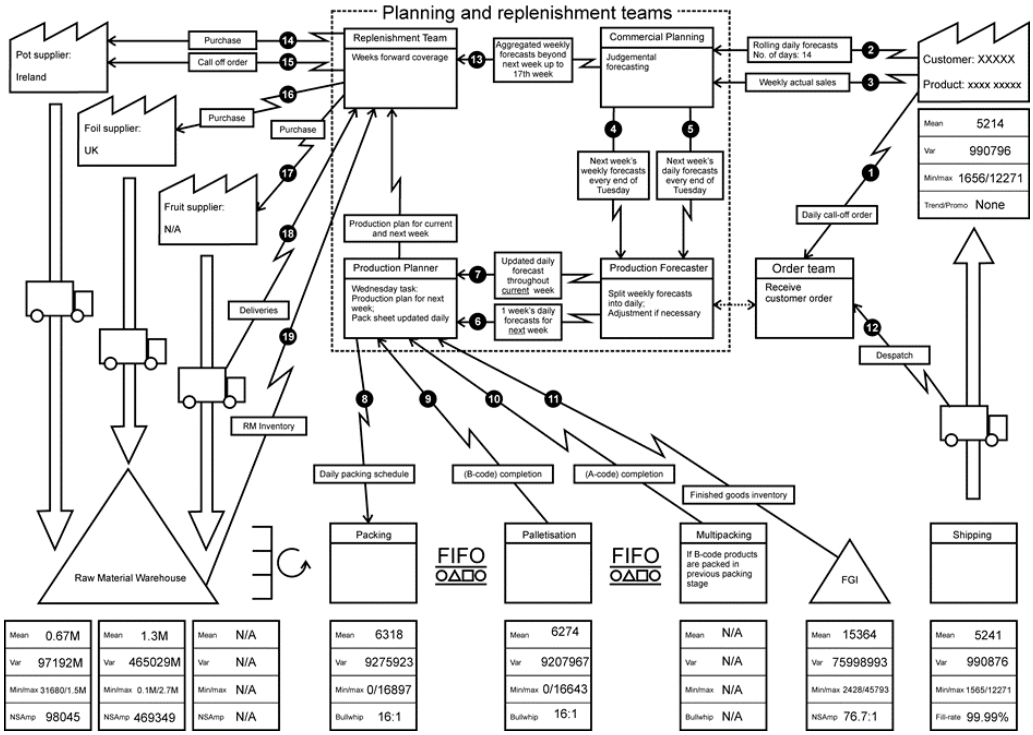

Figure 1 – Dynamic VSM

Demand forecasting

During the VSM development, we have observed that the first replenishment decision starts with customers' daily orders (1), rolling daily forecasts provided by customers (2) and actual weekly sales (3). This information is used to create weekly forecasts (4) that are then split into daily forecasts by allocating various percentages to each day (5). Daily forecasts are then passed to the production forecasting team for more adjustments (6, 7). For the chosen category of products, most of the weekly demand follows a stochastic pattern, without significant trend and seasonality (Figure 2a), except for those products which are on frequent promotions (Figure 2b). The daily demand exhibits a clear heartbeat-shaped, seasonality pattern throughout the week (Figure 2c).

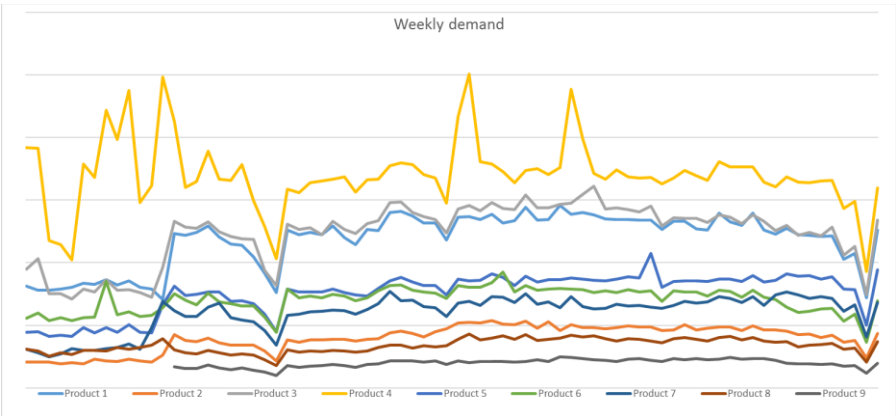

Figure 2a – Weekly demand for the 9 products

Figure 2b – Weekly demand of a highly promoted product

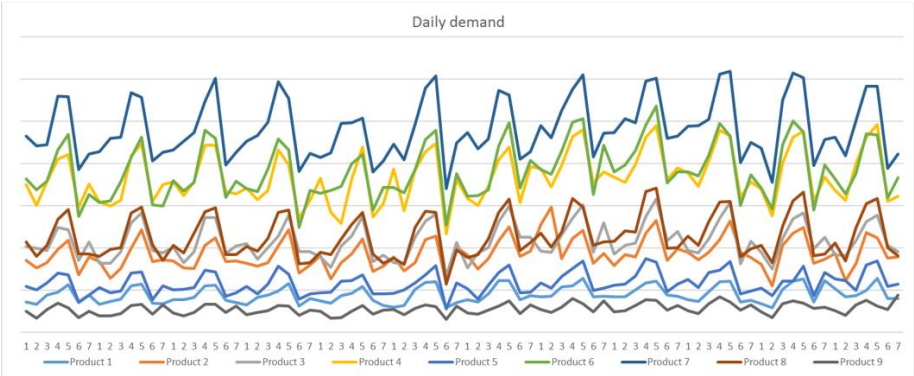

Figure 2c – Daily demand for the 9 products

By investigating the products' demand pattern and their forecasting process, we have found several issues. First, the company solely employs judgment to make forecasts, without any structure and guidance, and demand planners placed too much emphasis on forecast accuracy – resulting in a tendency to "keep chasing the noise". Fildes et al. (2009) observed similar behaviour, with demand planners making small forecast adjustments in order to gain accuracy but with opposite results. In general, where statistical forecasting methods such as simple exponential smoothing (SES) can easily achieve high accuracy for weekly demand forecasts at least (Hyndman and Athanasopoulos, 2013), judgmental forecasting can be costly and unnecessary. Table 1 shows a summary of the 9 products' MSEs for the company's judgmental forecasts and the forecasts given by its customers, versus a simple exponential smoothing with α=0.5 (with manual over-ride at Christmas, when judgemental forecasts perform best). There is a clear pattern that SES can outperform most of the current strategies employed.

Product	Exponential smoothing	Company	Customer	Best performer
Vanilla	12225	19300	35520	Exponential smoothing
Hazelnut	53319	68120	64489	Exponential smoothing
Raspberry	52583	78869	63907	Exponential smoothing
Lemon	120160	207338	140767	Exponential smoothing
Rhubarb	17660	45722	18389	Exponential smoothing
Strawberry	239079	271951	212784	Customer
Toffee	651575	712153	570016	Customer
Black Cherry	82991	134126	64041	Customer
Mango	3712	6733	12132	Exponential smoothing

Table 1 – A comparison of forecasting methods for the 9 products (MSE)

We also employed a measure of forecast nervousness (Li and Disney, 2017) to assess the accuracy of future forecasts used to order long lead-tile raw materials (such as plastic pots and labelling foils) and for financial planning. The variance of the j-step ahead order forecast error is:

$$\Delta[j] = \mathrm{var}\left(d_t - \hat{d}_{t-j|t}\right)$$

The demand in period *t* The forecast made at time *t-j*, of the order in time period *t*

As a forecast error in the near future is more costly (or at least harder to deal with) than one in the distant future, we adopt a geometrically weighted sum of order forecast error variances as a measure of nervousness:

$$Nervousness = \Delta = \sum_{j=1}^{\infty} w(1-w)^{j-1}\Delta[j]$$

Our analysis identified nervousness at w=0.8, company forecast nervousness at 12.89, naive forecast at 12.55 and exponential smoothing (α=0.5) at 8.88.

If we aggregate the forecasts into weekly buckets there is little to no seasonality and exponential smoothing with α= 0.5 performs really well, while nervousness is reduced by 30%. This further reinforces the fact that the forecasting team should be concerned with process supervision, not with chasing noise. It should also focus on generating forecasts once a week (i.e. on a Tuesday) rather than the current practice of drafting forecasts on Tuesdays, update them on Wednesday and finalise them on Thursday, with emergency revisions on Friday. The weekly forecasts should then be split into daily forecasts, for production purposes, rather than the current practice of using a rolling 2 week average to determine the daily split, which was introducing a lot of variation into the daily forecasts. A further recommendation is that the daily spilt should be based on a 3 month history of daily demands, updated every 3 months.

Time delay

The approach currently employed by the company results in a time-delay in the information flow. As the forecasting team is predicting next week's demand using the previous week's demand information, the production planning team uses relatively outdated information to produce weekly plans (8). The time-delay in the forecasting process is 14 days. Information and material flows in planning and production could be between 5-12 days long. It is argued that delays in information and material flow may contribute to bullwhip effect (Forrester 1961), although bullwhip can be eliminated by using alternative replenishment strategy when lagged demand is not available or used (Hosoda and Disney, 2012). In our current case, the latest demand information is simply not used in the process, whereas it is always recommended to take account of more updated information.

Ordering policy

Production - Currently the case company plans for weekly production within a bi-weekly cycle. They aim to produce each product three times a fortnight (which explains the two-week forecasting horizon). Production planning follows a strategy called "days forward coverage", that is, to produce the amount of products that cover certain following days' demand (according to the forecasts). They also pack two additional days' demand as safety stock. This set-up matches the staggered delivery system in Hedenstierna and Disney (2016). The 'days forward coverage' policy is commonly used in practice, but theoretically it is known to create large amounts of Bullwhip and it does not exploit inventory theory properly as it sets the safety stock to a function of average demand. Additionally, if the demand forecasts suffer from large

amounts of errors, the days forward coverage policy will potentially struggle to cope. The case company is also faced with production constraints, such as minimum and maximum production batch sizes due to large wastage costs, hence any established production policy (e.g. proportion-order-up-to policy) would be very challenging to implement here. Figure 3 bellow captures the trade-offs that the company has to reconcile when considering the packing frequency it adopts for each of the 9 case products.

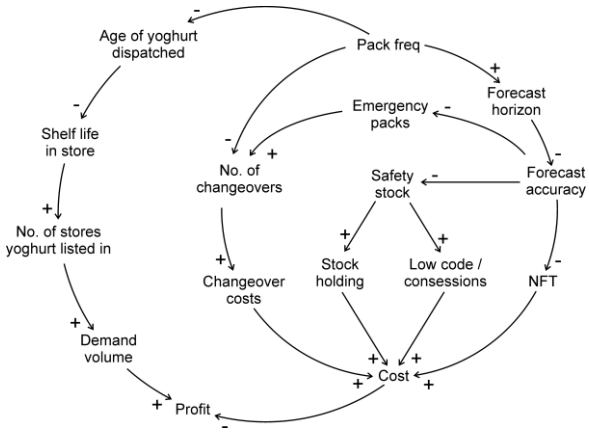

Figure 3 – Packing frequency trade-offs

Our first recommendation was to adopt a constant safety stock approach. Also known as a 'time-varying safety stock' Hedenstierna and Disney (2016) suggest this can be reviewed and updated periodically. Furthermore, we also need to consider the "*NFT*" situation – in a traditional ordering policy often the targeted finished goods inventory level is set as low as possible (or even zero). However, for the case company the opening stock should be enough to cover daily demand. When running a simulation for the 9 products (which follows the company's current policy but with fixed packing frequency and a fixed safety stock) the bullwhip was significantly reduced. Table 2 show the current level of bullwhip and simulated bullwhip.

Due to the high levels of bullwhip the company is confronted with, and less accurate forecasts, the NSAmp ratio for most of the 9 products is accordingly high (Table 2). However, in the simulation described above, the NSAmp is also reduced.

Product	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Bullwhip ratio comparison:									
Current system	21.47	17.70	26.24	43.89	33.34	47.92	19.60	18.04	38.56
Simulation	7.867	3.71	6.29	7.63	12.31	12.76	3.90	6.04	12.94
Finished goods NSAmp ratio comparison:									
Current system	31.348	28.286	38.886	46.970	37.239	53.393	36.506	32.194	36.949
Simulation	13.955	13.074	15.777	21.228	17.229	26.001	15.138	13.741	17.292

Table 2 – A comparison of Bullwhip and NSAmp ratios

Furthermore, as products are dispatched on a first-come-first-served basis, we were able to model the age, a_k of the product on the day it is dispatched to the customer.

$$a_1 = \begin{cases} 0, & p_t = 0, \\ \gamma, & p_t > 0, \end{cases} \quad \text{where } \gamma = \begin{cases} \min\{d_t - i_{t-1}, d_t\} & i_{t-1} < p_t \wedge d_t > i_{t-1}, \\ \min\{0, d_{t+1}\} & \text{otherwise.} \end{cases}$$

$$a_k = \begin{cases} 0, & p_t = 0, \\ \min\left\{d_{t+k-1}, p_t - \sum_{i=1}^{k-1} a_i, \max\left\{\sum_{i=0}^{k-1} d_{t+i} - i_{t-1}, 0\right\}\right\}, & p_t > 0. \end{cases}$$

Raw material replenishment - The raw material replenishment process starts with future weekly order streams for the next 17 weeks (13). These are used to create derived demand for raw materials that are used to create purchase orders to suppliers (14, 16, 18), and to call-off raw materials (15, 17, 19). The lead-time is supplier dependent (between 2 and 8 weeks). There are many different supplier replenishment loops, one for each raw material, resulting in many replenishment decisions. The case company is essentially ordering raw materials by using batched policy in order to achieve economies of scale. However, it is suggested that a smaller batch size will be more cost-effective and gain stability of orders (Holland and Sodhi, 2004), using for example a two-bin Kanban ordering system.

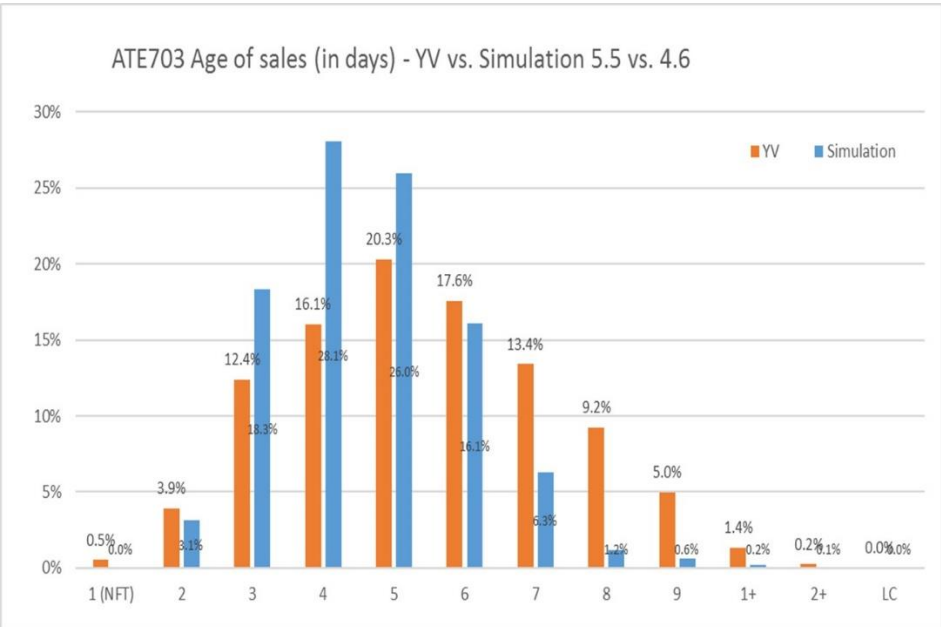

Figure 4 – Average age of products (Historical vs. Proposed performance)

Information sharing

There are multiple benefits of information sharing in order to reduce bullwhip in many aspects. Chen et al. (2000) suggested bullwhip can be reduced from sharing market demand from customer end. Demand information sharing is also recommended to reduce bullwhip in other literature (e.g. Lee et al., 1997). Wang and Disney (2016) summarised that information transparency and centralised decisions will be helpful to eliminate bullwhip. At the case company, the level of information transparency is relatively low, and the decision-making is also de-centralised. For market demand data, customers have not yet agreed to share market demand with the company.

The above findings support a series of improvement suggestions made to the case company, who is now trilling their implementations. These suggestions include the adoption of:

- Exponential smoothing forecasts (with $\alpha = 0.5$, no manual interventions);
- Daily production targets based on 3 month average forecasts;
- A constant safety stock and a days forward coverage linked to the duration till the next scheduled production;
- No NFT will be attempted;
- A kanban call-off system developed with key raw material suppliers.

Future work

While the results of our statistical analysis and simulation, combined with findings of the interviews and focus groups, support bullwhip effect theory, future work will focus on the development of a new model for the company's forecasting, planning and replenishment process(es). Our recommendations are now trialled to be implemented across the 9 case products, with each product being produced twice a week rather than thrice fortnightly. The aim is to have a production system with no NFTs, no low-code products and fresher products to be sold to the retailers, with the same (or better) financial performance. If successful, the company expects that the fresher products have the potential to increase the number of outlets they are sold at, leading to an increase in demand volumes and, ultimately, in increased profits. Future work also involves developing new working standards and changing the hearts and minds of people across the organisation in order to get a scalable and sustainable business improvement.

Conclusions

Our empirical case study examines the extent to which a company's operations and SC system impacts on the bullwhip effect, with a focus on demand forecasting, production planning and raw material replenishment processes. We also highlight multiple replenishment decisions and lead-times within a single echelon SC. In light of our findings, we suggest that the current stylised, theoretical SC models, and the knowledge derived from them, may need to be updated.

References

Cachon, P. C., Randall, T., and Schmidt, G. M. (2007), "In search of the Bullwhip effect", *Manufacturing & Service Operations Management*, 9(4), 457-479.

Cao, B., Xiao, Z., and Sun, J. (2016), "A study of the bullwhip effect in supply- and demand-driven supply chain", *Journal of Industrial and Production Engineering*, 34(2), 124-134.

Chen, F., Drezner, Z., Ryan, J.K., and Simchi-Levi, D. (2000), "Quantifying the bullwhip effect in a simple Supply chain: The impact of forecasting, lead times, and information", *Management Science*, 46(3), 436-443.

Croson, R., Donohue, K., Katok, E., and Sterman, J. (2013), "Order stability in supply chains: Coordination risk and the role of coordination stock", *Production and Operations Management*, 23(2), 176-196.

Disney, S.M., Gaalman, G., Hedenstierna, C.P.T., and Hosoda, T. (2015), "Fill rate in a periodic review order-up-to policy under auto-correlated normally distributed, possibly negative, demand", *International Journal of Production Economics*, 170, 501–512.

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

Fildes, R., Goodwin, F., Lawrence, F., and Nikolopoulos, K. (2009), "Effective forecasting and judgmental adjustments: an empirical evaluation and strategies for improvement in supply-chain planning", *International Journal of Forecasting*, 25(1), 3-23.

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

Ghali, M. A. (2003), "Production-planning horizon, production smoothing, and convexity of the cost functions", *International Journal of Production Economics*, 81, 67-74.

Hedenstierna, C.P.T. and Disney, S.M. (2016), "Inventory performance under staggered deliveries and auto-correlated demand". *European Journal of Operational Research*, 249 (3), 1082–1091.

Holland, W., and Sodhi, M.S. (2007), "Quantifying the effect of batch size and order errors on the bullwhip effect using simulation", *International Journal of Logistics Research and Applications*, 7(3), 251-261.

Hosoda, T., and Disney, S.M. (2012), "On the replenishment policy when the market demand information is lagged", *International Journal of Production Economics*, 135(1), 458-467

Hyndman, R.J. and Athanasopoulos, G. (2013), Forecasting: principles and practice. OTexts: Melbourne, Australia. http://otexts.org/fpp/. Accessed on <29/03/2018>.

Isaksson, O., Seifert, R. (2016), "Quantifying the bullwhip effect using two-echelon data: A cross-industry empirical investigation", *International Journal of Production Economics*, 171(3), 311-320.

Lee, H.L., Padmanabhan, P., and Whang, S. (1997), "Information distortion in a SC: the bullwhip effect", *Management Science*, 43, 543–558.

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

Li, Q. and Disney, S.M. (2016), "Revisiting rescheduling: MRP nervousness and the bullwhip effect". *International Journal of Production Research*, 55(7), 1992–2012.

Li, Q., Disney, S.M., and Gaalman, G. (2014), "Avoiding the bullwhip effect using Damped Trend forecasting and the Order-Up-To replenishment policy", *International Journal of Production Economics*, 149, 3-16

Naim, M.M., Dinsey, S.M. and Towill, D. (2004), "SC dynamics" in *Understanding SCs: Concepts, Critique and Futures*, Oxford University Press.

Ohno, T. (1988), Toyota Production System: Beyond Large-scale Production, Productivity Press, Portland, OR.

Rother, M. and Shook, J. (1999), *Learning to See*. Lean Enterprise Institute, Cambridge, MA, USA.

Sterman, J.D. (1989), "Modeling Managerial Behavior: Misperceptions of Feedback in a Dynamic Decision Making Experiment", *Management Science*, 35(3), 321-339.

Wackerly, D., Mendenhall, W., and Scheaffer, R.L. (2008), *Mathematical Statistics with Applications*, 7th Ed, Cengage Learning.

Wang, X., and Disney, S. M. (2015), "The bullwhip effect: Progress, trends and directions", European Journal of Operations Research, 250(2016), 691-701.