Ratanachote, P. and Disney, S.M., (2008), "On the square root law for bullwhip: The case of arbitrary lead-times and AR(1) demand", Fifteenth International Working Seminar on Production Economics, Innsbruck, Austria, March 3-7, pp199-212.

On the square root law for bullwhip: The case of arbitrary lead-times and AR(1) demand

Po-ngarm Ratanachote and Stephen M. Disney

Logistics System Dynamics Group, Cardiff Business School,
Colum Drive, Aberconway Building, Cardiff, CF10 3EU, UK.
RatanachoteP@Cardiff.ac.uk, DisneySM@cardiff.ac.uk

Abstract

We develop a model of a divergent distribution network with one distribution centre and *n* retailers. Each of the *n* retailers faces a stochastic first-order autoregressive demand. Furthermore, each retailer uses the Order-Up-To (OUT) policy with minimum mean squared error forecasting to generate replenishment orders that are placed onto the distribution centre (DC). The DC operates a base stock replenishment policy. The base stock policy is an OUT policy with time-invariant forecasts.

We assume that there are piece-wise linear and convex inventory holding and backlog costs. We are able to show that the "Square Root Law for Inventories" due to Maister [1] also holds when the OUT replenishment policy is present. Recall, that Maister's Square Root Law was derived for the Economic Lot Size model.

We assume that piece-wise linear and convex capacity costs are present. These costs are associated with over-time working (above an optimised production capacity) and the cost of un-used, or lost, capacity. These capacity costs have been added to capture the opportunity costs associated with the bullwhip effect. Under such a costing scheme we are also able to show that a "Square Root Law for Bullwhip" also exists when network consolidation occurs.

We also consider the impact of different lead-times at each retailer and the distribution centre, in both the de-centralised and centralised distribution network. Our analytical results that described the discrete time system behaviour are verified via a spreadsheet simulation model.

Keywords: Distribution network design, Square root law, Bullwhip, Lead-times

1. Introduction

Distribution Network Design (DND) is concerned with the placement of an arbitrary number of distribution centres (DC's) that act as stock holding facilities to enable the efficient flow of materials through a supply chain. The number of distribution centres is a decision variable; as is the size of the warehouse and their geographical position. These DND decisions are important as the location and number of DC's influence transportation costs and delivery / collection lead-times. The lead-times influence the amount of stock that must be held to provide a certain level of product availability. This in turn influences the capacity of the DC's that are required.

Observing a distribution network, one begins at the customer and then looks upstream through distribution centres and factories. Usually there exists a phenomenon of increased order variability as it proceeds up a supply chain. This is called the bullwhip effect. Lee et al. [2] and [3] identify four major causes of the bullwhip effect; demand forecast updating, order batching, price fluctuation, and rationing/shortage gaming. This study capture the relationship between distribution network designs and the cost associated with the bullwhip effect.

Our study is concerned with costs in both centralised and decentralised distribution networks. Two types of costs are considered; inventory related and capacity related costs. The inventory related costs includes inventory holding and backlog costs, while the capacity related costs associated with the bullwhip cost comprises of lost capacity and over-time costs. Both inventory and capacity costs are assumed to be piece-wise linear and convex. Analytical methods and a spreadsheet based simulation are used to investigate economic performance.

Our model is based on Disney, Saw and McCullen [4], where the square root law for bullwhip is found to exist in the case of independently and identically distributed (i.i.d) demand and unity lead-times. Here, we extend this finding for the case of first-order autoregressive demand and arbitrary lead-times at each retailer and the DC. We also assume that the Order-Up-To (OUT) replenishment policy with minimum mean squared error forecasting at the retailer's echelon and a base stock replenishment policy at the DC's echelon are assumed to exist. We believe our methodology to be unique in that it captures the link between the supply chain dynamics and the distribution network structure. At present these aspects do not appear to be incorporated into modern supply chain design software. Correctly accounting for these issues will be of much interest to industry companies with large distribution networks.

The structure of the paper is as follows. We review the literature and highlight Maister's "Square Roots Law for Inventory" in section 2. Section 3 defines the decentralised and centralised distribution networks and specifies the demand processes and cost structures that each player faces. Section 4 defines the replenishment rules used at the retailers and the DC's in the networks. In section 5 we derive the expressions for the variance of the order rates and inventory levels in the decentralised network that are required in the cost function analysis. Section 6 highlights the variances for the centralised distribution network. These two sets of results (for the decentralised and centralised distribution networks) are brought together in section 7 where we derive the "Square Root Law for Bullwhip". In section 8 we verify our results via simulation. Section 9 summaries the implications of this work and concludes.

2. Literature review

As noted above, our model is based on Disney, Saw and McCullen [4]. [4] and [13] provide some useful mathematical expressions of the optimal target net stock and the optimal capacity investment level. In addition, [4] and [13] provide expressions related to the inventory and capacity costs. In our study, we use a mixture of techniques including system engineering, statistical and simulation techniques. A block diagram has been used to present the flow and dynamic of the system as in [5], [6] and [7]. The block diagram also allows us to easily derive expressions of the system variances, which are important inputs for our analytical model.

The DND problem has been considered by many scholars. Chopra [8] gives the framework for designing a distribution network by concerning two dimensions: satisfy customers' needs and costs of doing so. In term of costs, Chopra [8] includes inventory and transportation costs as the main factors that affect the decision about the number of facilities required. According to Simchi-Levi et al. [9], there are many important issues associated with the management of a distribution network. For example, networks have to be configured, inventory controlled, transportation decisions made, fleets managed and truck routed. The distribution network problems are challenging by their need for system-wide cost minimisation and their uncertain nature. Supply chain dynamics is also mentioned as one of the causes that make these types of problems difficult to solve. Hammant et al. [10] present the use of a decision support system (DSS) of an automotive aftermarket supply chain. The service level and costs associated to distribution network design (inventory and transportation costs) are simultaneously considered. The results underline the benefit of network consolidation of the case study.

2.1. The square root law for inventory

Maister [1] introduced the "Square Root Law" for inventory costs when consolidation occurs in a distribution network. Quoting directly from Maister,

> "If the inventories of a single product (or stock keeping unit) are originally maintained at a number (*n*) of field locations (refereed to as the decentralised system) but are then consolidated into one central inventory (referred to as the centralised system), then the ratio
>
> $$\frac{\text{Decentralised system inventory}}{\text{Centralised system inventory}} = \sqrt{n} \tag{2.1}$$
>
> exists", Maister [1].

Maister [1] has provided a proof of the square root law for cycle stock under the assumptions that the Economic Order Quantity (EOQ) based on the Wilson Lot Size Formula is applied to control the inventory system.
In this study, we emphasise the behaviour of inventory and capacity costs associated with distribution network. The results from Maister [1] shows that the square root law is precisely represented the ratio of system inventories in the case that the demands in each decentralised location are identical. However, if the demand in each location is not homogeneous, the square root law is still a good approximation (also we provide an exact analytical expression for the components of the ratio but do not explicitly study it here). This means the consolidation distribution network requires less inventory and capacity which lead to reduced costs.

3. The analytical model

We will consider 2 different distribution network scenarios. A decentralised distribution network and a centralised distribution network. This will allow us to consider the impact the distribution network design on its dynamic and economic performance.

3.1. The decentralised distribution network

In this scenario, *n* customers each has a first-order autoregressive demand which is placed upon *n* retailers. Each retailer manages his inventory levels by placing orders onto a distribution centre. The orders are generated with the OUT policy. The policy exploits conditional expectation to generate MMSE forecasts of future demand, thus ensuring the retailers inventory costs are kept to a minimum. The *n* distribution centres manage their inventory with a base stock policy. Our decentralised distribution network is depicted in Figure 1.

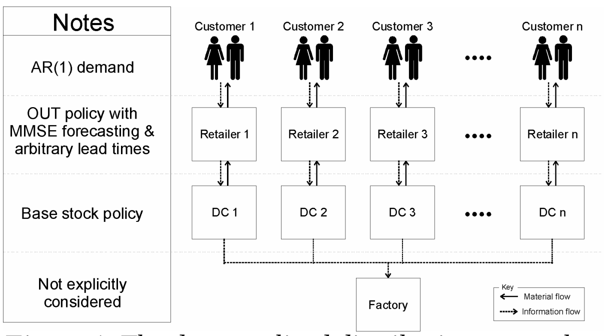

Figure 1. The decentralised distribution network

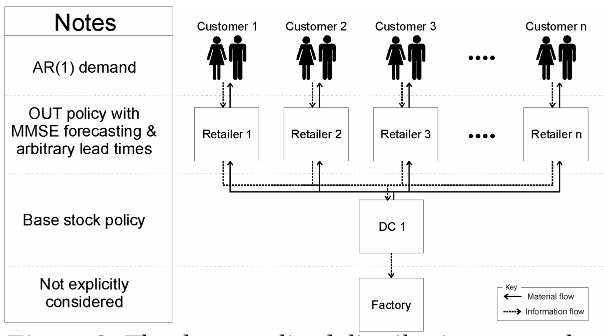

Figure 2. The decentralised distribution network

3.2. The centralised distribution network

In the centralised scenario, each of the *n* retailers faces a first-order autoregressive demand from its customer. Similarly, each retailer operates the customer order with the OUT policy with MMSE forecasting. However all of the retailers orders are aggregated as they are passed on to the DC. See Figure 2. The consolidated distribution centre generates its replenishment order with a base stock policy. The details of the replenishment policy of the decentralised and centralised systems are described in Section 4 and Section 5, respectively.

3.3. The demand model

We assume in both scenarios the demand that retailer $i,(i \in N^{+} < n)$ receives from his customer base is an Auto-Regressive random process of first order (AR(1)). Specifically we use the mean centred AR(1) demand process as we assume the mean is constant and known, see Equation (3.1).

$$\left.\begin{aligned} a_{i:t} &= f_i a_{i:t-1} + e_{i:t} \\ d_{i:t} &= a_{i:t} + m_{d:i} \end{aligned}\right\} \qquad (3.1)$$

In Equation (3.1), $d_{i:t}$ is the mean centred AR(1) demand at retailer i at time t. f_i is the auto-regressive constant of retailer i's demand, $|f_i| < 1$ and $a_{i:t}$ is a zero centred AR(1) process. $m_{d:i}$ is the mean demand at retailer i and $e_{i:t}$ is an identically and independently distributed random variable drawn from a normal distribution with a mean of zero and a standard deviation of $s_{e:i}$. We also assume that each retailers demand process is independent from other retailers demand processes, thus different e_i's, $s_{e:i}$ and $m_{d:i}$'s are independent from each other. $e_{i:t}$ is also independent from previous realisations of itself, $e_{i:t-x}$, $x \in N^+$. AR(1) processes have been found to represent a wide range of real life products, Lee So and Tang [11]. It is also well known that the long run variance of the AR(1) demand process is given by

$$\frac{s_{AR}^2}{s_e^2} = \frac{1}{1-f^2}. \qquad (3.2)$$

3.4. The costs in our distribution network

We assume, at all times, a linear system exists. Thus when the net stock falls below zero, we assume that demand is met from an alterative source and paid for with a premium. That is, the expediting strategy is used, as it is in Lee, So and Tang [11] and Gavirneni [12]. In these conditions it is usual to assume piecewise linear and convex inventory holding and backlog costs exist. That is costs are incurred in each period via the following equation.

$$\text{Inventory cost for period } t = \begin{cases} H(ns_t) & \text{, when } ns_t \geq 0, \\ B(-ns_t) & \text{, when } ns_t < 0. \end{cases} \qquad (3.3)$$

Thus H is the cost per period of holding one item of inventory and B is the cost, per period of having one unit of backlog. Figure 3 provides a visualisation of how these V-type inventory costs are incurred.

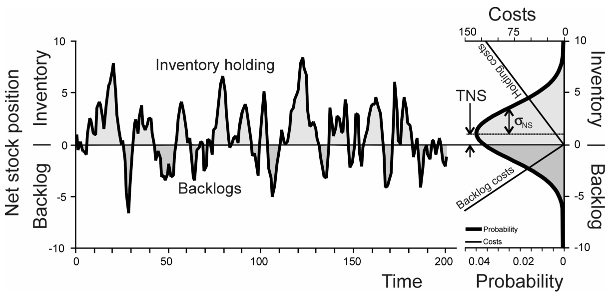

Figure 3. How inventory costs are generated over time

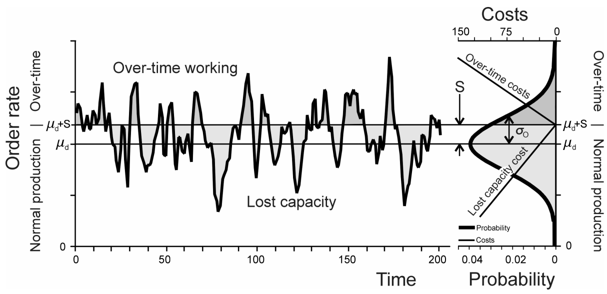

Figure 4. How the order costs are generated over time

As we assumed the error terms, $e_{i:t}$ are normally distributed, then the distribution of the in inventory levels (and the order rates which we will consider later on) will also be normally normal distributed. As such we can determine, via the probability density function, the optimal target net stock (TNS^*), in order to minimise the expected inventory holding and backlog costs. The TNS can be thought of as a safety stock. It is given by

$$TNS^* = s_{NS}\sqrt{2}\left(erf^{-1}\left[\frac{B-H}{B+H}\right]\right) = s_{NS}X_I \text{ with } X_I = \sqrt{2}\left(erf^{-1}\left[\frac{B-H}{B+H}\right]\right) \qquad (3.4)$$

where s_{NS} is the standard deviation of the net stock levels and erf^{-1} is the inverse error function, a function that is related to the normal distribution. Disney et al [13] provides more details on the derivation of Equation (3.4). We remark that by setting the target net stock to Equation (3.4) will result in a critical fractile of periods ending in a positive net stock position as is commonly found in the newsvendor problem. Thus we may use H and B to ensure a strategy level of inventory availability, the P1 service measure, Silver, Pyke and Peterson [14]. When the target net stock to set to TNS^* then we also remark that the inventory cost per period is given by Equation (3.5). Here we can see that the inventory costs are a linear function of the standard deviation of the net stock levels over time.

$$I_{£} = \frac{s_{NS}(B+H)e^{-erf^{-1}\left[\frac{B-H}{B+H}\right]^2}}{\sqrt{2p}} = s_{NS}Y_I \text{ with } Y_I = \frac{(B+H)e^{-erf^{-1}\left[\frac{B-H}{B+H}\right]^2}}{\sqrt{2p}}. \quad (3.5)$$

In a similar manner we may assume that capacity costs are also piecewise linear and convex. Consider Figure 4. If the order rate is above a certain periods normal capacity (of $m_d + S$) then over-time work is used to produce those exceptionally high orders. If over-time is not available then this proportion of the replenishment orders could be sourced from a subcontractor at a premium price. However, if the available capacity is not used in a particular period, then it is assumed that an opportunity cost is incurred as the investment in production capacity is not being fully exploited. Thus the following equation describes the capacity cost incurred in each period.

$$\text{Capacity cost for period } t = \begin{cases} N(S+m_d-o_t) & \text{, when } o_t \le (S+m_d) \\ P(o_t-(S+m_d)) & \text{, when } o_t > (S+m_d) \end{cases} \quad (3.6)$$

These capacity costs are identical in nature to the inventory costs. N is the cost of not fully exploiting the available capacity, and thus $N(S+m_d-o_t)$ is the cost of lost capacity in each period. P is the premium paid for overtime (or subcontracting) working to produce (or distribute) over the available capacity level. Thus $P(o_t-(S+m_d))$ is the over-time cost incurred in each period. Figure 4 conceptualises this point.

It is easy to see that there is an optimal amount of capacity to invest in, S^*, just as there was an optimal amount of safety stock (TNS). The optimal capacity level above (or below) the mean demand is given by

$$S^* = s_O\sqrt{2}\left(erf^{-1}\left[\frac{P-N}{P+N}\right]\right) = s_O X_O \text{ with } X_O = \sqrt{2}\left(erf^{-1}\left[\frac{P-N}{P+N}\right]\right) \quad (3.7)$$

where s_O is the standard deviation of the order rate. In a similar manner to the inventory costs, the following expression for the capacity costs can be derived,

$$C_{£} = \frac{s_O(N+P)e^{-erf^{-1}\left[\frac{P-N}{P+N}\right]}}{\sqrt{2p}} = s_O Y_O \text{ with } Y_O = \frac{(N+P)e^{-erf^{-1}\left[\frac{P-N}{P+N}\right]}}{\sqrt{2p}}. \quad (3.8)$$

Disney et al [13] explores this further if interested readers would like more detail. Later we will assume that the unit costs are H=1, B=9, N=4 and P=6 are present at all locations in the distribution network.

4. The replenishment decisions in the distribution network

There are two different types of replenishment decisions in our distribution network. The retailers use an "order-up-to" policy (OUT) with minimum mean squared error (MMSE) forecasting. The MMSE forecast over the arbitrary (but known and constant) lead-time (and review period) is given by conditional expectation.

The distribution centres however, will use the base stock policy. This is an OUT policy with a constant forecast of future demand. This is a simple policy that is mathematically tractable, especially in the light of the fact that it is rather tedious (but not impossible) to obtain the conditional expectation of demands from many customers with different demand processes that have passed though retailers with different lead-times. However, this complexity is avoided with the base stock policy as the base stock model is simply the OUT policy with constant forecasts. Given that we know the mean demand from all of the customers, we know the mean demand from all the retailers, and their combination. Thus, it is easy to set the appropriate order-up-to level in the base stock policy.

4.1. Retailer's replenishment policy: OUT policy with MMSE forecasting

Lets consider the replenishment policy used at each of the retailers. As stated before it is an OUT policy with conditional expectation forecasting, reacting to AR(1) customer demand. In an OUT policy the replenishment orders at retailer *i*, at time *t* are given by

$$o_{i:t} = \hat{d}_{i:t} - ns_{i:t} - wip_{i:t} \tag{4.1}$$

where $\hat{d}_{i:t}$ is the conditional expectation of demand over the lead-time and review period. For the AR(1) demand process over a lead-time of Tp periods (Note: Tp does not include the review period) then this forecast is given by

$$\hat{d}_{i:t} = a_{i:t}\left(\frac{f_i\left(f_i^{1+Tp}-1\right)}{f_i-1}\right)+m_{d:i}\left(1+Tp_i\right). \tag{4.2}$$

The net stock evolves by the usual inventory balance equation, see Equation (4.3). Here we can also see the influence of the known, constant, but arbitrary lead-time, $Tp \in N_0$. Recall the demand rate was defined earlier in Equation (3.1).

$$ns_{i:t} = ns_{i:t-1} + o_{i:t-(Tp_i+1)} - d_{i:t} \tag{4.3}$$

Finally we need an expression for the work in progress (WIP), $wip_{i:t}$, in Equation (4.1). This is given by

$$wip_{i:t} = \sum_{j=1}^{Tp} o_{i:t-j} \tag{4.4}$$

which we can see is simply the accumulation of the last Tp orders.

4.2. The distribution centre's replenishment rule: Base stock policy

The replenishment rule at the distribution centre is given by the base stock policy. A base stock policy is an OUT policy with constant, time invariant forecasts. In fact the forecasts have been set to be equal to the long run average of the total demand DC *i* faces, $m_{d:i}^{DC}$. Thus the order rate in the base stock policy is given by

$$o_{i:t}^{DC} = m_{d:i}^{DC}\left(1+Mp_i\right) - ns_{i:t}^{DC} - wip_{i:t}^{DC} \tag{4.5}$$

where the superscript (*DC*) has been used to denote that here the orders (*o*), mean demand (m_d), net stock (*ns*) and WIP (*wip*) refer to the i^{th} DC. The subscript *t* again refers to time. In fact the only difference between (4.5) and (4.1) is that the forecasted demand at the DC level is simply the mean aggregated demand (multiplied by Mp+1) the particular DC faces.

The inventory balance at each DC has conceptually remained the same. We only need to add the superscript (*DC*) to highlight the fact we are considering DC *i*. We have also used the term Mp (rather than Tp) for the distributor's replenishment lead-time, see Equation (4.6).

$$ns_{i:t}^{DC} = ns_{i:t-1}^{DC} + o_{i:t-(Mp_i+1)}^{DC} - d_{i:t}^{DC} \tag{4.6}$$

These changes are also applied to the WIP equation, see Equation (4.7).

$$wip_{i:t}^{DC} = \sum_{j=1}^{Mp} o_{i:t-j}^{DC} \tag{4.7}$$

5. The decentralised distribution network scenario

We will first characterise the variances of the orders in the decentralised network. We will then consider the economic performance of the decentralised distribution network.

5.1. Retailer variance analysis

In order to obtain expression for the variances in the system it is convenient to first express the replenishment decisions as a block diagram. This block diagram can then be investigated to yield the necessary variance expressions.The block diagram for the retailer's replenishment decisions is shown in Figure 5. Here you can see we have represented the policy using *z*-transforms as is common in discrete control theory. It is a simple task for a skilled control engineer to construct this block diagram directly from the difference equations highlighted in Equations (4.1)-(4.4). We refer interested readers to a good control engineering text book for more information on this aspect of our study. We recommend Nise [15].

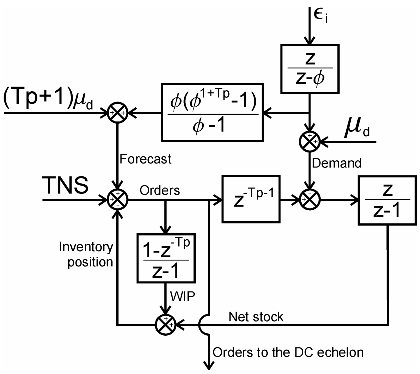

Figure 5. Block diagram of a retailer's replenishment decision: OUT policy with MMSE forecasting

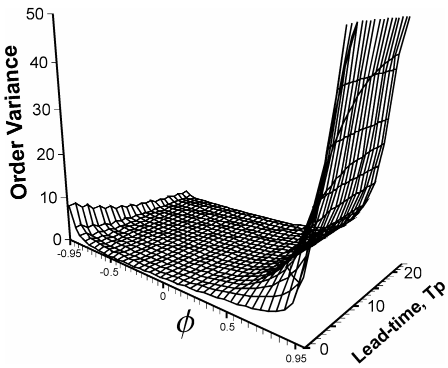

Figure 6. The behaviour of the variance of the order rate (Note: This also holds for the distribution centre, but we have not yet discovered this)

Arranging the block diagram using standard techniques (Nise, [15]) results in the following transfer function that relates the error term to the order rate. It is

$$\frac{O_i(z)}{e_i(z)} = \frac{f_i - z + (z-1) f_i^{2+Tp_i}}{(z - f_i)(f_i - 1)} . \tag{5.1}$$

Notice that we have now included the subscript i to denote the fact that this is the transfer function for the i'th retailer. This was omitted from Figure 5 for clarity. Taking the inverse *z*-transform of (5.1) results in the time domain impulse response for the order rate. This is

$$o_{i:t} = Z^{-1}\left[\frac{O_i(z)}{e_i(z)}\right] = \frac{f_i^t \left(f_i^{2+Tp_i} - 1 - \left(f_i^{1+Tp_i} - 1\right) h[t-1]\right)}{f_i - 1} \tag{5.2}$$

where $h[x]$ is the Heaviside step function, that is $h[x]$=1 if $x \ge 0$, 0 otherwise. From Equation (5.2) we can sum its square to determine the long run variance of the order rate at a retailer. This operation is known as Tsypkin's Relation, Tsypkin [16]. We refer to Disney and Towill [17] for more information.

$$s_{o:i}^2 = \sum_{t=0}^{\infty} o_{i:t}^2 = \frac{1 + f_i + 2 f_i^{4+2Tp_i} - 2 f_i^{2+Tp_i}(1+f_i)}{s_{e:i}^{-2}(f_i - 1)^2 (1+f_i)} \tag{5.3}$$

This variance expression holds for each of the i=1 to n retailers. Figure 6 detais Equation (5.3). Now let's consider the net stock levels. The transfer function of interest is

$$\frac{NS_i(z)}{e_i(z)} = \frac{z^{-Tp_i}\left(z - f + z^{2+Tp_i}(f_i - 1)_i + f_i^{2+Tp_i}(1-z)\right)}{(z-1)(f_i - 1)(f_i - z)} . \tag{5.4}$$

The time domain impulse response for this system is given by the inverse z-transform of Equation (5.4). It is

$$ns_{i:t} = Z^{-1}\left[\frac{NS_i(z)}{e_i(z)}\right] = \frac{f_i^{1+t}-1}{1-f_i}h[Tp_i - t]. \tag{5.5}$$

From here we use Tsypkin's Relation again to reveal the variance of the net stock. It is

$$s^2_{ns:i} = \sum_{t=0}^{Tp} ns^2_{i:t} = \frac{f_i^{4+2Tp_i} - 1 + 2f_i(1+f_i) - 2f_i^{2+Tp_i}(1+f_i) + Tp_i(f_i^2-1)}{s^{-2}_{e:i}(f_i-1)^3(1+f_i)}. \tag{5.6}$$

Notice that here we have exploited our knowledge of the fact that the net stock is zero when *t>Tp*, is the summation.

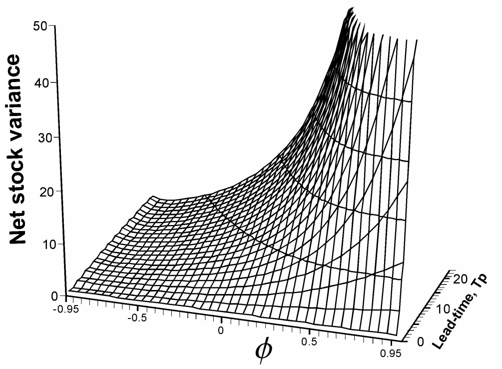

Figure 7. The behaviour of the variance of the retailer's net stock

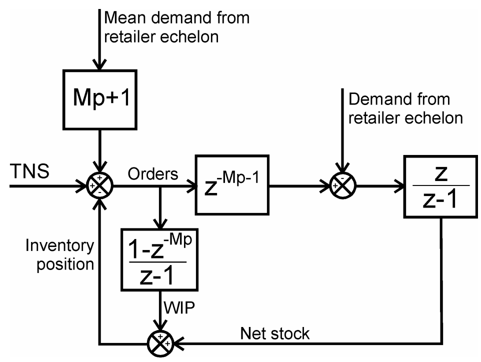

Figure 8. Block diagram of a DC's replenishment decision: The base stock policy

5.2. Distribution centre variance analysis

Figure 8 below highlights the block diagram of the DC's replenishment decision. In order to be read properly it needs to be coupled with the block diagram of the retailer. We have, however, not joined them directly here, as leaving them separated allow us to use the same two block diagrams for both the centralised and decentralised scenarios.

Amazingly the transfer function that relates the order rate at a DC to a particular customers error term (e_i) is given by

$$\frac{O_i^{DC}(z)}{e_i(z)} = \frac{O_i(z)}{e_i(z)}. \tag{5.7}$$

Thus the variance of order rate at DC *i* is the same as the order rate at retailer *i*. This means that the variance of the order rate at each DC *i* is simply

$$s^2_{DC\,o:i} = \frac{1+f_i+2f_i^{4+2Tp_i} - 2f_i^{2+Tp_i}(1+f_i)}{s^{-2}_{e:i}(f_i-1)^2(f_i+1)}. \tag{5.8}$$

This is logical as the base stock policy simply "passes on orders". Thus, Figure 6 also illustrates the DC's order variance. However, there has been a subtle change in the way the net stock behaves that we need to work out explicitly. The transfer function that relates the i^{th} error term to the i^{th} DC is given by

$$\frac{NS_i^{DC}(z)}{e_i(z)} = \frac{z^{-Mp_i}\left(z^{1+Mp_i}-1\right)\left(f_i - z + (z-1)f_i^{2+Tp_i}\right)}{(1-z)(z-f_i)(f_i-1)}. \tag{5.9}$$

Taking the inverse z-transform yields

$$ns_{i:t}^{DC} = Z^{-1}\left[\frac{NS_i^{DC}(z)}{e_i(z)}\right] = \frac{\begin{pmatrix} -(f_i-1)f_i^{Mp_i}\left(f_i^{2+t+Tp_i}-1\right)- \\ \left(f_i^t - f_i^{1+Mp_i}\right)\left(f_i^{1+Tp_i}-1\right)h[t-Mp_i-1]- \\ \left(f_i^{Mp_i}-f_i^t\right)\left(f_i^{2+Tp_i}-1\right)h[t-Mp_i] \end{pmatrix}}{f_i^{Mp_i}\left(f_i-1\right)^2}. \tag{5.10}$$

Finally summing the squared time domain impulse response yields the inventory variance at each DC i. It is given by

Here we can see that the variance of the net stock at the DC is always greater than the variance of the net stock at an equivalent retailer. This is due the fact that MMSE forecasting is not used at the DC. We refer interested readers to Hosoda and Disney [7] for more discussion of this aspect.

$$s_{DC:nsi}^2 = s_{e:i}^2\sum_{t=0}^{\infty}\left(ns_{i:t}^{DC}\right)^2 = s_{e:i}^2\left(\sum_{t=0}^{Tp}\left(\frac{f_i^{1+t}-1}{1-f_i}\right)^2+\sum_{t=Tp+1}^{\infty}\left(ns_{i:t}^{DC}\right)^2\right) = s_{nsi}^2 + s_{e:i}^2\sum_{t=Tp+1}^{\infty}\left(ns_{i:t}^{DC}\right)^2$$

$$= \frac{\begin{pmatrix} f_i^{2Mp_i}\begin{pmatrix} f_i^2+Mp_i\left(f_i^2-1\right)-1+ \\ f_i^{2+Tp_i}\left(f_i^{1+Mp_i}-1\right)\left(f_i^{2+Tp_i}-2-2f_i+f_i^{3+Mp_i+Tp_i}\right)\end{pmatrix}- \\ f_i^{4+2Tp_i}\left(f_i^2\right)^{Mp_i}\left(f_i^{1+Mp_i}-1\right)^2 \end{pmatrix}}{s_{e:i}^{-2}f_i^{2Mp_i}(f_i-1)^3(f_i+1)} \tag{5.11}$$

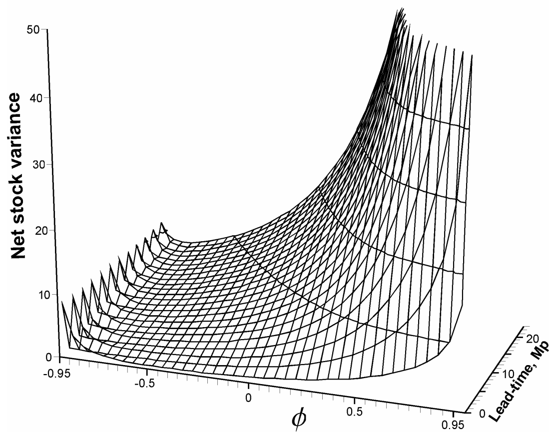

Figure 9. The variance of the DC's net stock levels when the retailer's lead-time is unity (Tp=1)

5.3. Cost analysis in the decentralised distribution centres

In the decentralised distribution network the here are n DC's. The optimal amount of safety stock at DC i is given by

$$TNS_i^* = X_I\sqrt{s_{DC:nsi}^2} \tag{5.12}$$

where X_I is the function related to the backlog and holding costs given in Equation (3.4) and $s_{DC:nsi}^2$ is the variance of the net stock levels at DC i given in Equation (5.11). When the target net stock is set in this way then the inventory cost at each DC i is given by

$$I_{£:i} = Y_I\sqrt{s_{DC:nsi}^2} \tag{5.13}$$

where Y_I is the function related the inventory holding and backlog costs highlighted in Equation (3.5). Thus the total inventory related cost across all DC's in the decentralised distribution network is given by

$$I_£ = Y_I\sum_{i=1}^{n}\sqrt{s_{DC:nsi}^2}. \tag{5.14}$$

If all the demand processes are the same for each customer (that is, if $f_i = f$ and $s^2_{e:i} = s^2_e$) and all the lead-times at the retailers are identical (that is, if $Tp_i = Tp$) and all the lead-times at the n distribution centres are the same (that is, if $Mp_i = Mp$) then the inventory costs in the distribution centre are

$$I_£ = nY_I\sqrt{s^2_{DCns}}\,. \qquad (5.15)$$

In Equation (5.15) the subscript i has been dropped as all the customers, retailers and distribution centres are the same. Consider now the order (capacity) related costs. The optimal amount of slack capacity (above the average demand, $m_{d:i}$) at DC i is given by

$$S^*_i = X_O\sqrt{s^2_{DC\alpha i}} \qquad (5.16)$$

where X_O is the function related to the lost capacity and overtime costs given in Equation (3.7) and $s^2_{DC\alpha i}$ is the variance of the order rate at DC i. Similarly for the capacity related costs at DC i we have

$$C_{£:i} = Y_O\sqrt{s^2_{DC\alpha i}}\,. \qquad (5.17)$$

Thus the total capacity costs in all of the DC's is

$$C_£ = Y_O\sum_{i=1}^{n}\sqrt{s^2_{DC\alpha i}}\,. \qquad (5.18)$$

If all the demand processes are the same for each customer (that is, if $f_i = f$ and $s^2_{e:i} = s^2_e$) and all the lead-times at the retailers are identical (that is, if $Tp_i = Tp$) then the capacity costs in the distribution centre are then

$$C_£ = nY_O\sqrt{s^2_{DCo}}\,. \qquad (5.19)$$

Again in Equation (5.19) the subscript i has been dropped as all the customers, retailers and distribution centres are the same.

6. The centralised distribution network scenario

In the centralised distribution there are still i=1 to n customers and i=1 to n retailers. However, there is only one distribution centre. Thus, all of the retailer's orders are placed onto the one DC.

6.1. Retailer variance analysis in the centralised network

As nothing has changed in either the customer or the retailer echelon of the supply chain then the variance of retailer i's order rate or net stock levels remains unchanged from the previous section. Thus, Equation (5.3) characterises retailer i's order rate variance and Equation (5.6) describes retailer i's net stock variance. Of course Figures 6 and 7 also illustrate the variances at the retailer in the centralised network.

6.2. DC variance analysis in the centralised network

As there is only one DC then the variance of the order rate at the DC is equal to the sum of all of the variances from each of the inputs into the system (that is, the sources of noise, $e_{i:t}$). Thus, the variance of the order rate at the single DC in the centralised DC network is given by

$$s^2_{C:DCo} = \sum_{i=1}^{n}\frac{1+f_i+2f_i^{4+2Tp_i}-2f_i^{2+Tp_i}(1+f_i)}{s^{-2}_{e:i}(f_i-1)^2(f_i+1)}\,. \qquad (6.1)$$

Here we have pre-pended the subscript with a C to denote the fact that we are considering the centralised distribution network. In a similar manner the variance of the single DC's inventory level is given by

$$s^2_{CDCns} = \sum_{i=1}^{n} \frac{\begin{pmatrix} f_i^{2Mp}\left(f_i^2 + Mp\left(f_i^2-1\right)-1+f_i^{2+Tp_i}\left(f_i^{1+Mp}-1\right)\left(f_i^{2+Tp_i}-2-2f_i+f_i^{3+Mp+Tp_i}\right)\right)- \\ f_i^{4+2Tp_i}\left(f_i^2\right)^{Mp}\left(f_i^{1+Mp}-1\right)^2 \end{pmatrix}}{s_{e:i}^{-2} f_i^{2Mp}(f_i-1)^3(f_i+1)}. \tag{6.2}$$

Notice in Equations (6.1) and (6.2) that there is no need for a subscript to *Mp* as there is only one DC.

If we consider that all the demand processes are the same for each customer (that is, if $f_i = f$ and $s^2_{e:i} = s^2_e$) and all the lead-time at the retailers are identical (that is, if $Tp_i = Tp$) then the DC's order and net stock variances reduce down to Equations (6.3) and (6.4) respectively.

$$s^2_{C:DC:o} = n\left(\frac{1+f+2f^{4+2Tp}-2f^{2+Tp}(1+f)}{s_e^{-2}(f-1)^2(f+1)}\right) = ns^2_{DC:o} \tag{6.3}$$

$$\begin{aligned} s^2_{CDCns} &= n\frac{\begin{pmatrix} f^{2Mp}\left(f^2+Mp\left(f^2-1\right)-1+f^{2+Tp}\left(f^{1+Mp}-1\right)\left(f^{2+Tp}-2-2f+f^{3+Mp+Tp}\right)\right)- \\ f^{4+2Tp}\left(f^2\right)^{Mp}\left(f^{1+Mp}-1\right)^2 \end{pmatrix}}{s_e^{-2} f^{2Mp}(f-1)^3(f+1)} \\ &= ns^2_{DC:ns} \end{aligned} \tag{6.4}$$

6.3. DC cost analysis in the centralised network

It is easy to see now that the safety stock at the DC is given by TNS^*,

$$TNS_i^* = X_I\sqrt{s^2_{C:DCns}} \tag{6.5}$$

and the inventory related costs at the single DC are

$$I_{£} = Y_I\sqrt{s^2_{C:DCns}}\ . \tag{6.6}$$

The optimal amount of capacity at the DC in the centralised scenario is

$$S_C^* = X_O\sqrt{s^2_{C:DC:o}} \tag{6.7}$$

and the capacity related costs are

$$C_{£:C} = Y_O\sqrt{s^2_{C:DC:o}}\ . \tag{6.8}$$

7. The square root law for bullwhip

We are now ready to reveal the "Square Root Law for Bullwhip". It holds under the conditions that all the demand processes have the same structure at each customer (that is, if $f_i = f$ and $s^2_{e:i} = s^2_e$) and all the lead-times at the retailers are identical (that is, if $Tp_i = Tp$). Under these conditions the capacity related costs in the decentralised scenario are equal to

$$C_{£} = nY_O\sqrt{\frac{1+f+2f^{4+2Tp}-2f^{2+Tp}(1+f)}{s_e^{-2}(f-1)^2(f+1)}} \tag{6.9}$$

and the capacity related costs in the centralised scenario are

$$C_{£:C} = Y_O\sqrt{n\left(\frac{1+f+2f^{4+2Tp}-2f^{2+Tp}(1+f)}{s_e^{-2}(f-1)^2(f+1)}\right)}. \tag{6.10}$$

Dividing the Equation (6.9) by (6.10) reveals the "Square Root Law for Bullwhip"

$$\frac{\text{Decentralised capacity cost}}{\text{Centralised capacity cost}} = \frac{nY_O\sqrt{s^2_{DC:o}}}{Y_O\sqrt{ns^2_{DC:o}}} = \sqrt{n}\ . \tag{6.11}$$

It is also interesting to note that the slack capacity (that is the capacity investment above the average demand) also behaves with a square root law. Explicitly we have

$$\frac{\text{Decentralised slack capacity}}{\text{Centralised slack capacity}} = \frac{nX_O\sqrt{\sigma^2_{DC:o}}}{X_O\sqrt{n\sigma^2_{DC:o}}} = \sqrt{n}\,. \tag{6.12}$$

We remind readers that if these assumptions (identical demand parameters and identical retailer lead-times) do not hold, the exact ratios can be determined form the expressions provided in sections 5 and 6. We are also able to verify that the "Square Root Law for Inventory" also holds in our scenario with OUT replenishment policies, for both the inventory cost and safety stock requirements. Recall that Maister [1] defined it for the Economic Lot Size model. This is because

$$\frac{X_I\sqrt{\sigma^2_{DC:ns}}}{X_I\sqrt{\sigma^2_{CDC:ns}}} = \frac{Y_I\sqrt{\sigma^2_{DC:ns}}}{Y_I\sqrt{\sigma^2_{CDC:ns}}} = \sqrt{n}\,. \tag{6.13}$$

8. Verification via simulation

In this section, the results from the previous sections are validated using a spreadsheet model. To verify that the square root law for bullwhip, we simulate two distribution networks; decentralised and centralised. We assume here that there are n=2 customers / retailers, and the lead-times in both retailer and DC's echelon are unity, Tp=1 and Mp=1 (although our analytical results hold for all lead-times). Also assume that the unit costs are H=1, B=9, N=4 and P=6 are present at all locations in the distribution network. Table 1 shows the results from the simulation, which included different settings of the autoregressive constants (ϕ). For each case of the autoregressive constant, we compare inventory and capacity cost between theoretical formula (the upper figure as derived in Sections 5.3 and 6.3) with the results from simulation (the lower figure). The last column shows the ratio between capacity costs in decentralised and centralised networks. All ratios are equivalent to 1.414, which is the squared root of two ($\sqrt{n}$, where here n=2). Therefore our simulation verifies the square root law for bullwhip which we revealed in Section 7. In fact, the square root law also holds for the cases of inventory costs. Furthermore, our simulation has extended to the case that n is greater than 2 with arbitrary Tp_i and Mp_i. The results confirm that the square root law is robust for the cases with large distribution networks and different lead-times at each retailers and DC's.

ϕ	Retail echelon			
	TNS^*	S^*	Inventory cost	Capacity cost
0.95	2.80874	1.003080	3.84598	15.2965 (theory)
			3.84815	15.2985 (simulation)
0.5	2.31035	0.444863	3.16384	6.78395
			3.16325	6.78482
0	1.81239	0.253347	2.48192	3.86343
			2.48102	3.86325
-0.5	1.43282	0.193497	1.96213	2.95074
			1.96225	2.95225
-0.95	1.28315	0.736306	1.75718	11.2283
			1.75759	11.2252

Table 1. Verification of the "Square Root Law for Bullwhip" via simulation and theory
(Table 1 is continued overleaf)

9. Managerial implications and concluding remarks

We have presented a methodology to dynamically design a distribution network based on inventory and capacity costs. At each point in the network the optimum safety stock has been set to minimise the sum of the inventory holding and backlog costs in the face of the stochastic demand pattern that each location faces. We have also defined the optimum capacity at each point in the network in order to minimise the sum of the lost capacity and over-time working at each location. The maths that we have exploited to achieve this is based on a linear system, thus inventory has been backlogged rather then lost when a negative inventory position has occurred, and production over the capacity limit has been made up in over-time (or provided via a sub-contractor with the same lead-time).

We have shown that the capacity (or bullwhip) costs behave in exactly the same way as the inventory related costs when distribution networks are consolidated. We have achieved this in a supply chain with AR(1) demands, arbitrary lead times and the OUT replenishment policy. This is an extension of Disney, McCullen and Saw, [4] where only i.i.d. demands and unity lead-times were considered.

f	Distribution echelon								Ratio (1)/(2)
	Decentralised network				Centralised network				
	TNS^*	S^*	Inv cost	Capacity cost (1)	TNS^*	S^*	Inventory cost	Capacity cost (2)	
0.95	9.11331	1.003108	24.9599	30.593	12.8882	1.41857	17.6493	21.6325	$\sqrt{n}$
			24.9551	30.597			17.6492	21.6325	1.414
0.5	3.2986	0.444863	9.03433	13.5679	4.66492	0.629131	6.38824	9.59396	$\sqrt{n}$
			9.03263	13.5696			6.38876	9.59498	1.414
0	1.81239	0.253347	4.96385	7.72685	2.56311	0.358287	3.50997	5.46371	$\sqrt{n}$
			4.96204	7.72649			3.50882	5.46496	1.414
-0.5	1.25457	0.193497	3.43606	5.90148	1.77423	0.273646	2.42966	4.17298	$\sqrt{n}$
			3.43699	5.90353			2.42861	4.17449	1.414
-0.95	1.2393	0.736306	3.39425	22.4567	1.75264	1.04129	2.4001	15.8793	$\sqrt{n}$
			3.39539	22.4523			2.4001	15.8768	1.414
Notes	At each DC		In the DC echelon		At the single DC		In the DC echelon		

Table 1 (continued). Verification of the “Square Root Law for Bullwhip” via simulation and theory

The “Square Root Law for Bullwhip” suggests that reasons to consolidate distribution networks are actually a lot stronger than previously thought. The likely impact of this is to force companies to consolidate even further than they have in the past, increasing the amount of traffic on the road. Thus, internalising the external costs transportation causes is now even more important and as good guardians of the environment we should accept higher fuel prices and / or taxes.

Further work could include extending our approach to include multiple products and coupling our methodology with readily available commercial software to determine the optimal placement of facilities via “centre of gravity” modelling. At present our methodology is only able to evaluate and compare a rather small range of logistical scenarios. This may be useful for companies who have limited range of options available for the structure of their distribution network. However, more complex network structures could also be considered. Aspects such as cross-docking, inter DC trunking and more general flows (rather than just divergent flows) could be considered.

10. Acknowledgements

We would like to thank the Commission on Higher Education, Thailand, for the financial support on Ms. Ratanachote's PhD study.

11. References

1. Maister, D.H., 1976. Centralisation of Inventories and the 'Square Root Law'. *International Journal of Physical Distribution*, 6 (3), 124-134.
2. Lee, H.L., Padmanabhan, P., Whang, S., 1997. The bullwhip effect in supply chains. *Sloan Management Review*, 38 (3), 93-102.
3. Lee, H.L., Padmanabhan, P., Whang, S., 1997. Information distortion in a supply chain: the Bullwhip Effect. *Management Science*, 43, 543-558.
4. Disney, S.M., McCullen, P., Saw, R.J., 2006. Dynamic supply chain design: square root law for Bullwhip. *Proceedings of the International Symposium on Logistics*, Beijing, China, July 9th -12th.
5. Chen, Y.F., Disney, S.M., 2006. The myopic order-up-to policy with a proportional feedback controller. *International Journal of Production Research*. Vol. 45, No. 2, pp351-368.
6. Disney, S.M., Towill, D.R., Van de Velde, W., 2004. Variance amplification and the golden ratio in production and inventory control. *International Journal of production Economics*, 90(3), 295-309.
7. Hosoda, T. and Disney, S.M., 2006. On variance amplification in a three-echelon supply chain with minimum mean squared error forecasting. *OMEGA: The International Journal of Management Science*, 34, 344-358.
8. Chopra, Sunil, 2003. Designing the distribution network in a supply chain. *Transportation Research Part E*, 39, 123-140.
9. Simchi-Levi, D., Chen, X., Bramel, J., 2005. *The Logic of Logistics*, Springer, New York.
10. Hammant, J., Disney, S.M., Childerhouse, P., Naim, M.M., 1998. Modelling the consequences of a strategic supply chain initiative of an automotive aftermarket operation. *International Journal of Physical Distribution & Logistics Management*, 29(9), 535-550.
11. Lee, H.L., So, K.C. and Tang, C.S. 2000. The value of information sharing in a two level supply chain. *Management Science*, 46 (5), 543-558.
12. Gavirneni, S., 2006. Price fluctuations, information sharing and supply chain performance. *European Journal of Operational Research*, 174, 1651-1663.
13. Disney, S.M., Chen, Y.F., Gaalman, G., Lambrecht, M., Towill, D.R., Van de Velde, W., Warburton, R.D.H., 2006. Costing the bullwhip effect, *INFORMS Annual Conference,* Hong Kong, June 25th-27th.
14. Silver, E.A., Pyke, D.F. and Peterson, R. (1985), "Inventory management and production planning and scheduling", *John Wiley and Sons*, New York.
15. Nise, N.S. 1995. Control systems engineering. *The Benjamin Cummings Publishing Company, Inc.*, Reading.
16. Tsypkin, Y.Z., 1964. Sampling systems theory and its application. *Vol. 2, Pergamon Press,* Oxford.
17. Disney, S.M., Towill, D.R., 2003. Vendor-managed inventory and bullwhip reduction in a two-level supply chain. *International Journal of Operations & Production Management*, 23(6), 625-651.
18. Box, G.E. and Jenkins, G.M., 1970. Time series analysis forecasting and control. *Holden-Day,* San Francisco.