

Impact of scheduling frequency and shared capacity on production and inventory costs

Carl Philip T. Hedenstierna¹, Stephen M. Disney²

Logistics Systems Dynamics Group, Cardiff Business School, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK. Tel: +44(0)2920 876310. Fax: +44(0)2920 874301.

¹ cpthed@gmail.com, ² DisneySM@cardiff.ac.uk.

Abstract

A multi-product production planning problem is considered. Products can either share the same production line, or be produced on different lines. We study the influence of planning cycle length on inventory and production performance, first in terms of how these affect the variability of the states in the production system, and then how this translates into cost. Production costs are assumed to be dominated by labour costs, where we use a model with guaranteed hours and overtime. Inventory performance is measured using holding and backlog costs. The replenishment policy used is the Proportional Order-Up-To (POUT) policy, containing as a special case the classical OUT policy. It does this by having a controllable feedback parameter, which alters the relative dynamics of inventory and capacity. Demand is assumed to be normally, independently and identically distributed. Our analytical results show that there exists an optimal rescheduling frequency that minimises total cost. This frequency is determined by the relationship between the costs and lead time. The economic benefit from planning in larger buckets follows from ‘demand pooling’, which allows more effective use of guaranteed hours and reduces overtime. A balance must be struck between reducing inventory costs, which generally get smaller as we plan more frequently, and labour costs, which increase as we plan more frequently. We also demonstrate that the pursuit of ever-faster rescheduling leads to severe cost penalties. A comparison between the benefits of optimising the rescheduling frequency, production line consolidation, and POUT parameter tuning highlights the economic potential of optimising rescheduling frequency.

Keywords: Rescheduling frequency, OUT policy, Inventory and capacity costs, multiple products

1. Introduction

Production plans are typically made on a cyclical basis, where production orders for the next period are issued, taking into account demand, forecasted demand, work-in-process levels, as well as the need for maintaining inventory around target levels. From the basic Order-Up-To (OUT) policy to more sophisticated applications of Manufacturing Resource Planning (MRP II), a rescheduling frequency must be set. This paper investigates the impact the rescheduling frequency when production costs and inventory costs are considered when production is scheduled with the Proportional Order-Up-To (POUT) policy.

Alfred P. Sloan could be one of the first promoters of high-frequency rescheduling. In 1924, Sloan decided that General Motors (GM) Corporation’ production should be rescheduled every ten days instead of once every three months as was done previously. During the 1920’s GM managed to increase its total inventory turnover from 2 to nearly 7½ times p.a., due to various efforts of streamlining GM’s production and distribution network (Sloan, 1963). Fast rescheduling was also considered important at Toyota Motor Corporation. Ohno (1988) describes that slow rescheduling of production causes large inventory swings if the market demand should change unexpectedly, in turn leading to fluctuating production rates (Bicheno and Holweg, 2009). Toyota used a planning cycle of ten days, and considered moving to weekly, or even daily cycles (Shingo, 1989). The importance of fast rescheduling was also noted by John Burbidge, who states it as one of his “five golden rules to avoid bankruptcy” (Burbidge, 1983). Fast rescheduling was implemented within the Period Batch Control

method, which was used for the manufacture of the Spitfire aircraft. (Burbidge, 1988) states that the rescheduling frequency should be as high as capacity permits, and that set-up time reduction allows for further shortening of the scheduling cycle.

Contrary to the idea of fast rescheduling is the concept of variability pooling, which occurs when two statistical populations are combined. As the variances are added, the resulting standard deviation per unit will be smaller. This logic is commonly used when production lines, or stock keeping units are consolidated (Maister, 1976). A common assumption is that faster rescheduling cycles offer a better trade-off between capacity requirements and inventory levels. This paper seeks to investigate the mechanism underlying this belief, as well as to quantify the relative benefits of some popular rescheduling frequencies.

We first present our model in Section 2. We then evaluate the consequences of various rescheduling frequencies in Section 3. Section 4 investigates the economics of altering the rescheduling frequency compared with other production system improvement strategies. Finally, we summarise our findings and conclude in Section 5.

2. Model development

Daily demand for product $k \in \{1..n\}$ is assumed to be an independent and identically distributed random variable drawn from a normal distribution

$$d_{k,t} \in N(\mu_k, \sigma_k). \quad (1)$$

When the daily demand is aggregated into planning cycles of P days long the demand during each planning cycle is given by

$$d_{P,k,t} \in N(\mu_k P, \sigma_k \sqrt{P}). \quad (2)$$

Notice how we have used the 3-tuple in the subscripts. The first space is for, $P \in \mathbb{N}$, the length of the planning cycle, dimensioned in days. The second space is for the product index k , where $k = 1$ to n products. The third space is for time period t . Demand is satisfied from inventory. The inventory balance equation is given by

$$ns_{k,t} = ns_{k,t-1} + o_{k,t-T_p-1} - d_{k,t}, \quad (3)$$

where $i_{k,t}$ is the inventory of product k at time t , $o_{k,t}$ is the production orders for product k started at time t that will be completed a lead-time T_p and review period later. The lead-time, $T_p \in \mathbb{N}_0$, is dimensioned in number of planning periods. The real-world lead-time (not including the sequence of events delay), $L \in \mathbb{R} > 0$, is dimensioned in days. The relationship between $\{T_p, L, P\}$ is given by $T_p = \lfloor \frac{L}{P} \rfloor$. We generate production order via the so-called Proportional Order-Up-To (POUT) policy [8]. The POUT orders are given by

$$o_{k,t} = \underbrace{\hat{d}_{k,t,t+T_p+1}}_{\text{Forecast}} + f_k \left(\underbrace{TNS + \sum_{j=1}^{T_p} \hat{d}_{k,t,t+j}}_{DWIP} - \left(\underbrace{i_{k,t-1} + \sum_{m=1}^{T_p} o_{k,t-m}}_{WIP} \right) \right). \quad (4)$$

Here $\hat{d}_{k,t,t+T_p+1}$ is the forecasted demand for product k , made at time t of demand in the period $T_p + 1$. f_k is the proportional feedback controller used for damping the WIP and inventory feedback information for product k . $0 \leq f_k < 2$ is required for stability. When $f = 1$ then the classical Order-Up-To S policy is present. S is the order-up-level and is made up of the sum of the safety stock or Target Net Stock (TNS) and the Desired Work In Progress ($DWIP$) or target "orders placed but not yet received". $DWIP$ is the sum of the forecasted demand over the lead-time, T_p . From the order-up-to level we subtract the inventory position. The inventory position is made up of the net stock for product k in time t , $i_{k,t}$ and the actual Work In Progress, $WIP_{k,t}$. $WIP_{k,t}$ is the previous T_p orders for product k , that have been placed, but have not yet been received. The sequence of events within the planning cycle is illustrated in Figure 1.

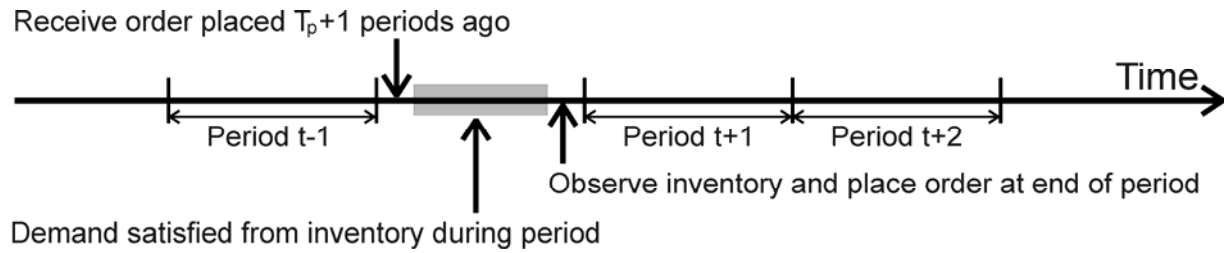


Figure 1. Sequence of events.

As we have assumed that the daily demand is i.i.d. normally distributed then optimal forecasts, with minimum mean squared error, are simply multiples of the average demand. The multiples that are required are related to the lead-time, T_p and the length of the planning cycle, P and we require the following two forecasts within the POUT policy,

$$\left. \begin{aligned} \hat{d}_{k,t,t+j} &= P\mu_k \\ DWIP_k &= \sum_j^{T_p} \hat{d}_{k,t,t+j} = T_p P\mu_k \end{aligned} \right\} \quad (5)$$

Notice that in order to ensure linearity we have allowed the net stock position to become negative. A negative net stock position means that a backlog is present. Customers are willing to wait for their product, but we penalise backlogs with a cost per unit backlogged per day, B . Inventory holding is also penalised via inventory holding cost per unit per day, H . Thus the following inventory cost function holds

$$\text{Inventory cost in period } t \text{ for product } k = P \left(H_k (ns_{k,t})^+ + B_k (-ns_{k,t})^+ \right), \quad (6)$$

where $(x)^+$ is the maximum operator, that is $(x)^+ = \max[0, x]$. Note that when we move from planning every day to buckets of P days then we need to scale the inventory holding and backlog costs by P . The Target Net Stock (TNS^*) level, is given by

$$TNS_k^* = \sigma_{NS,k} z_k \quad (7)$$

which will ensure that $\frac{B_k}{B_k + H_k} \times 100\%$ of periods end with inventory in stock (the availability metric). $1 - \frac{B_k}{B_k + H_k} = \frac{B_k + H_k}{B_k}$ is known as the "economic stock out probability" as a critical

proportion of periods end in stock-out. This critical proportion minimises the expected inventory cost over time. Here is the standard deviation of the net stock levels over time and

$$z_k = \Phi^{-1} \left[\frac{B_k}{B_k + H_k} \right], \quad (8)$$

where $\Phi^{-1}[x] = \sqrt{2} \operatorname{erf}^{-1}[2x - 1]$ is the inverse of the cumulative normal distribution function. When the *TNS* is set to this optimal level for each of the $k = 1$ to n products, the annual inventory related cost (we assume that 5 days \times 4 weeks \times 12 months = 240 days make up a year) will be given by

$$I_{\text{f}} = \sum_{k=1}^n 240 \sigma_{ns,k} (B_k + H_k) \varphi[z_k] \quad (9)$$

where $\varphi[x] = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]$ is the probability density function of the standard normal distribution (Disney et al., 2012). We notice in (9) that the expected inventory holding and backlog costs are linearly related to the standard deviation of the net stock levels $\sigma_{ns,k}$. It should be remembered that this requires that the safety stock, the Target Net Stock (*TNS*) has been set according to (8). If this is not the case then the inventory costs are not linear in $\sigma_{ns,k}$ and (9) does not hold.

For capacity costs we consider a situation with guaranteed hours and overtime, detailed as Case 3 in Disney et al. (2012), another paper in these conference proceedings. In every period we pay a standard rate of U_k per unit produced within a capacity level of $K_k + \mu_k$, no matter the actual production quantity. If production quantities are greater than the normal capacity, an overtime cost per unit, W_k , applies for all orders exceeding normal capacity. This is given by

$$\text{Capacity cost for period } t = U_k (K_k + \mu_k) + W_k (o_{k,t} - (K_k + \mu_k))^+ \quad (10)$$

where $K_k^* = \sigma_o z$ is the optimum slack capacity, which minimises the sum of normal production cost and the overtime cost. In K_k^* , $z_k = \Phi^{-1} \left[\frac{W_k - U_k}{W_k} \right]$, gives the optimum proportion of normal to overtime hours, while σ_o is the standard deviation of the production orders. When an optimal slack capacity is present, the annual capacity costs are

$$C_{\text{f}}^* = \sum_{k=1}^n 240 \left(\frac{W_k \sigma_{o,k}}{P} \varphi \left[\Phi^{-1} \left[\frac{W_k - U_k}{W_k} \right] \right] + U_k \mu_k \right). \quad (11)$$

In (11) we note that each product is effectively produced with its own separate capacity / workforce. If the capacity / workforce is flexible (in that it can produce any product with equal ability) then the orders can be combined and the capacity costs will be given by

$$C_{\text{f},J}^* = 240 \left(\frac{W_k \sqrt{\sum_{k=1}^n P \sigma_{o,k}^2}}{P} \varphi \left[\Phi^{-1} \left[\frac{W_k - U_k}{W_k} \right] \right] + \sum_{k=1}^n U_k \mu_k \right). \quad (12)$$

The cost functions contain expressions for the standard deviation of net stock (for inventory cost) and for the standard deviation of orders (for capacity cost). To obtain these, we must know the variance of demand and the variance amplification caused by the replenishment

policy. The variance ratios for the POUT policy have been derived previously in (Disney et al., 2004). (13) describes the order rate variance and (14) the inventory variance. These expressions assume that we are using a minimum mean squared error forecast, where the forecast equals average long-term demand.

$$\sigma_{o,k} = \sqrt{\frac{Pf_k\sigma_k^2}{2-f_k}} \quad (13)$$

$$\sigma_{ns,k} = \sqrt{\sigma_k^2 \left(P \left(1 + \left\lfloor \frac{L}{P} \right\rfloor + \frac{(f_k-1)^2}{f_k(2-f_k)} \right) \right)} \quad (14)$$

Figure 2 illustrates how the variance of the inventory and orders behave for all lead-times (L) and planning cycle lengths, P . We note that the order variance is independent of the lead-time L , but the inventory variance is affected by L . The rescheduling frequency affects both the order variance and the inventory variance via P . The order variance is an increasing and convex function in f , which is zero at the origin. The inventory variance is a convex function in f with a minimum at $f=1$. The sum of the order and inventory variance results in a convex function with a single minimum between $0 \leq f \leq 1$.

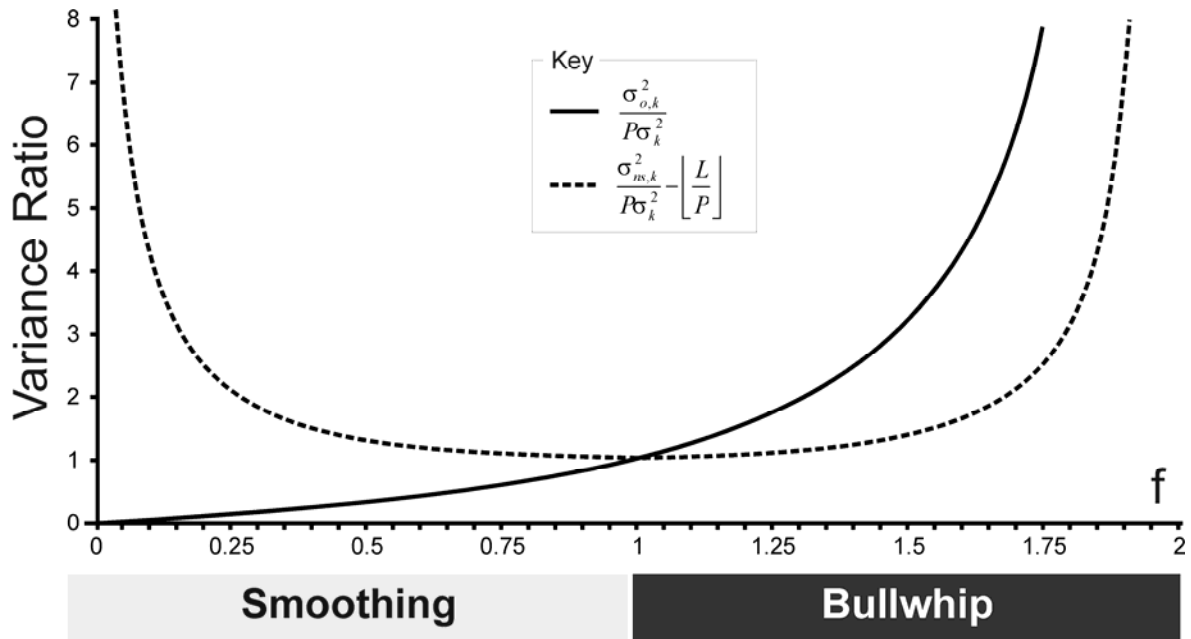


Figure 2. Variance of the orders and inventory generated by the POUT policy.

3. Optimising the length of the planning cycle

3.1. The OUT policy with individual production lines

Let's consider a two product case with the OUT policy, (hence $f=1$) and separate production lines. The total costs in this scenario are obtained by bringing together equations (9) (11), (13) and (14) to create a cost function as shown in (15). Note that we have assumed that policy factors such as P and f , the costs, B , H , U and W and the demand parameters μ and

σ , are the same for each product and that we have been dropped the subscripts (k) in order to save space.

$$C = \left(\begin{array}{c} \text{Capacity costs} \\ 240 \left(2\mu U + \frac{w}{P} \sqrt{\frac{P\sigma^2}{2\pi}} \exp \left[-\text{erf}^{-1} \left[\frac{w-2U}{W} \right]^2 \right] \right) + \\ \text{Inventory costs} \\ 240(B+H) \sqrt{\frac{2}{\pi}} \sqrt{P\sigma^2 (1+T_p)} \exp \left[-\text{erf}^{-1} \left[\frac{B-H}{B+H} \right]^2 \right] \end{array} \right) : T_p = \lfloor \frac{L}{P} \rfloor \quad (15)$$

The factor that complicates matters is that C is a non-linear function in P due to the floor function that is used to obtain the lead-time expressed in units of the replenishment cycle $T_p = \lfloor \frac{L}{P} \rfloor$. This is needed to correctly determine the net stock variance and the inventory costs. Let us first explore (15) numerically, see Figure 3. Here we have plotted C^* with dots for different P . We see that the dots have discontinuous breaks in their progression and these occur at the divisors of L , $P|L$ for even L and at the divisors of $L+1$, $P|L-1$ for odd L . We can see when $P > L$ then $T_p = 0$ and the solutions for P always lie on the cost line given by $C|_{T_p=0}$. Whenever we pass through a divisor of L , $P|L$, to the left, then the lead-time, expressed as in units of the number of planning cycles, T_p increases by one. The solutions when $P \leq L$ lie between two curves. The curves are given by $C|_{T_p=P}$ and $C|_{T_p=\lfloor \frac{L}{P} \rfloor + 1}$. The first curve will go through the divisors of L at $P = P|L$ ($P = P|L - 1$). All of the other solutions will lie below $C|_{T_p=P}$ but above $C|_{T_p=\lfloor \frac{L}{P} \rfloor + 1}$. For any given P , $C|_{T_p=0} < C|_{T_p=P-\lfloor \frac{L}{2} \rfloor - 1} < C|_{T_p=P}$. Furthermore due to (15) we know that $C|_{P=\lfloor \frac{L}{y} \rfloor + 1} \leq C|_{P=\lfloor \frac{L}{y+1} \rfloor + 1}$ because

$$(L+1)\sigma^2 (1+T_p) \leq (\lfloor \frac{L}{2} \rfloor + 1)\sigma^2 (2+T_p) \quad (16)$$

and

$$\frac{w}{L+1} \sqrt{(L+1)\sigma^2 / 2\pi} < \frac{w}{\lfloor L/2 \rfloor + 1} \sqrt{(\lfloor L/2 \rfloor + 1)\sigma^2 / 2\pi} . \quad (17)$$

This means that the optimal length of the planning cycle, P^* , will always come from a solution that lies on the $C|_{T_p=0}$ line. As we can see from Figure 3 $C|_{T_p=0}$ has a minimum in it. If the minimum lies below $P = L+1$ then we should set $P^* = L+1$. If the minimum lies above $L+1$ then we should set P^* to be the minimum.

We can find the minimum in $C|_{T_p=0}$ line by differentiating (15) with $T_p = 0$ w.r.t. P

$$\frac{dC}{dP} = \frac{120}{(P\sigma^2)^{3/2}} \sqrt{\frac{2}{\pi}} \sigma^2 \left(\exp \left[-\text{erf}^{-1} \left[\frac{B-H}{B+H} \right]^2 \right] (B+H) P \sigma^2 - \exp \left[-\text{erf}^{-1} \left[\frac{w-2U}{W} \right]^2 \right] W \sigma^2 \right), \quad (18)$$

from which the following first order conditions can be obtained,

$$x^* = \frac{w}{B+H} \exp \left[\text{erf}^{-1} \left[\frac{2B}{B+H} - 1 \right]^2 - \text{erf}^{-1} \left[1 - \frac{2U}{W} \right]^2 \right] = \frac{w}{B+H} \exp \left[\left(\Phi^{-1} \left[\frac{B}{B+H} \right] / \sqrt{2} \right)^2 - \left(\Phi^{-1} \left[1 - \frac{U}{W} \right] / \sqrt{2} \right)^2 \right] \quad (19)$$

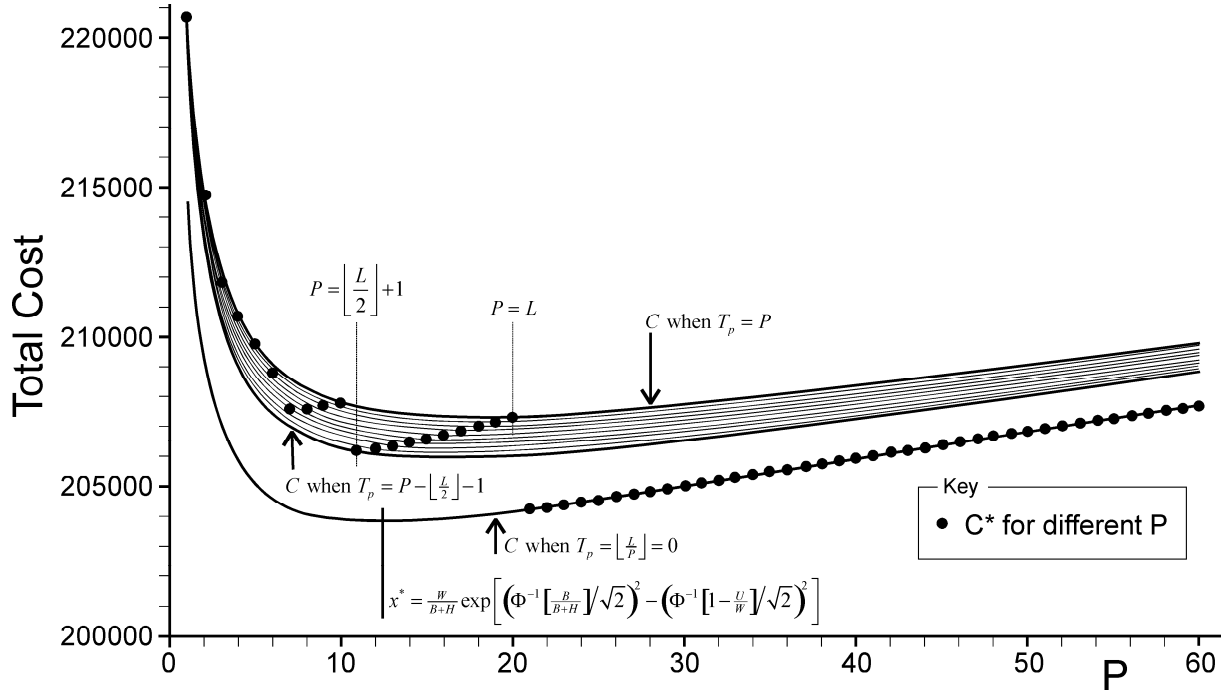


Figure 3. Total cost for the OUT policy with respect to P .

It is interesting to note that x^* is solely a function of the costs involved rather than the lead-time. Bringing (19) together with our insights from Figure 3 together we obtain (20) that describes the optimal length of the planning cycle P^* when we have 2 separate production lines.

$$P^* = (x^*, L+1)^+ ; x^* = \frac{w}{B+H} \exp \left[\left(\Phi^{-1} \left[\frac{B}{B+H} \right] / \sqrt{2} \right)^2 - \left(\Phi^{-1} \left[1 - \frac{U}{w} \right] / \sqrt{2} \right)^2 \right]. \quad (20)$$

3.2. The OUT policy with shared production capacity

Let's now consider the case of the OUT policy when the two products share production capacity. That is, they are both produced on the same production line. The cost function to minimise can be made up from (9), (12), (13) and (14),

$$C = 240 \left(\begin{aligned} & 2\mu U + \frac{W\sigma^2}{\sqrt{\pi}\sqrt{P}\sigma^2} \exp \left[-\text{erf}^{-1} \left[\frac{w-2U}{w} \right]^2 \right] + \\ & (B+H) \sqrt{\frac{2}{\pi}} \sqrt{P}\sigma^2 \left(1 + \left\lfloor \frac{L}{P} \right\rfloor \right) \exp \left[-\text{erf}^{-1} \left[\frac{B-H}{B+H} \right]^2 \right] \end{aligned} \right). \quad (21)$$

By following similar reasoning as above we are able to show that when we have shared production capacity the following expression can be obtained that describes the optimal length of the planning cycle.

$$P^* = (x^*, L+1)^+ ; x^* = \frac{w}{\sqrt{2}(B+H)} \exp \left[\left(\Phi^{-1} \left[\frac{B}{B+H} \right] / \sqrt{2} \right)^2 - \left(\Phi^{-1} \left[1 - \frac{U}{w} \right] / \sqrt{2} \right)^2 \right]. \quad (22)$$

The only difference between (22) and (20) is the square root in the denominator, hinting at the possibility that a “square root” law [7] exist for an arbitrary number of products.

3.3. The POUT policy with individual production lines

The case of the proportional order-up-to policy is a little harder to study than the classical OUT policy. The aspect that complicates matters is the fact that we have to optimise the feedback parameter f for each planning period P , before we can determine the total costs. The total costs are given by

$$C = 240 \left(2 \left(\mu U + \frac{W}{P} \sqrt{\frac{fP\sigma^2}{2\pi(2-f)}} \exp \left[-\operatorname{erf} \left[\frac{W-2U}{W} \right]^2 \right] \right) + \right. \\ \left. (B+H) \sqrt{\frac{2}{\pi}} \sqrt{P\sigma^2 \left(\frac{1}{2f(1-f)} + \left\lfloor \frac{L}{P} \right\rfloor \right)} \exp \left[-\operatorname{erf} \left[\frac{B-H}{B+H} \right]^2 \right] \right). \quad (23)$$

The derivative of (23) w.r.t. f is

$$\frac{dC}{df} = \frac{240\sigma}{(f-2)^2} \sqrt{\frac{2}{\pi}} \left(\frac{W}{\sqrt{\frac{fP\sigma}{2-f}}} \exp \left[-\operatorname{erf}^{-1} \left[\frac{W-2U}{W} \right]^2 \right] + \frac{P(f-1)(B+H)}{f^2 \sqrt{P\sigma \left(\frac{1}{f(2-f)} + \left\lfloor \frac{L}{P} \right\rfloor \right)}} \exp \left[-\operatorname{erf}^{-1} \left[\frac{B-H}{B+H} \right]^2 \right] \right) \quad (24)$$

Unfortunately (24) is a fourth order expression in f , and while its solution is obtainable, it is very lengthy, so we have not detailed it. We are able to plot it, see

Figure 4, which shows how the optimal f changes with L for the following numerical scenario: demand for each product has a mean of 10 and a standard deviation of 2; the inventory holding cost, $H = \text{£}1$ per day; the inventory backlog cost $B = \text{£}9$ per day and product; the unit cost production cost with normal working hours is $U = \text{£}40$ per unit; and the unit cost in over-time working is $W = \text{£}60$; the real-world lead-time is $L = 20$. It is interesting to note that the optimal f^* is never unity (so the OUT policy is never optimal), but f^* approaches unity as the lead-time $L \rightarrow \infty$.

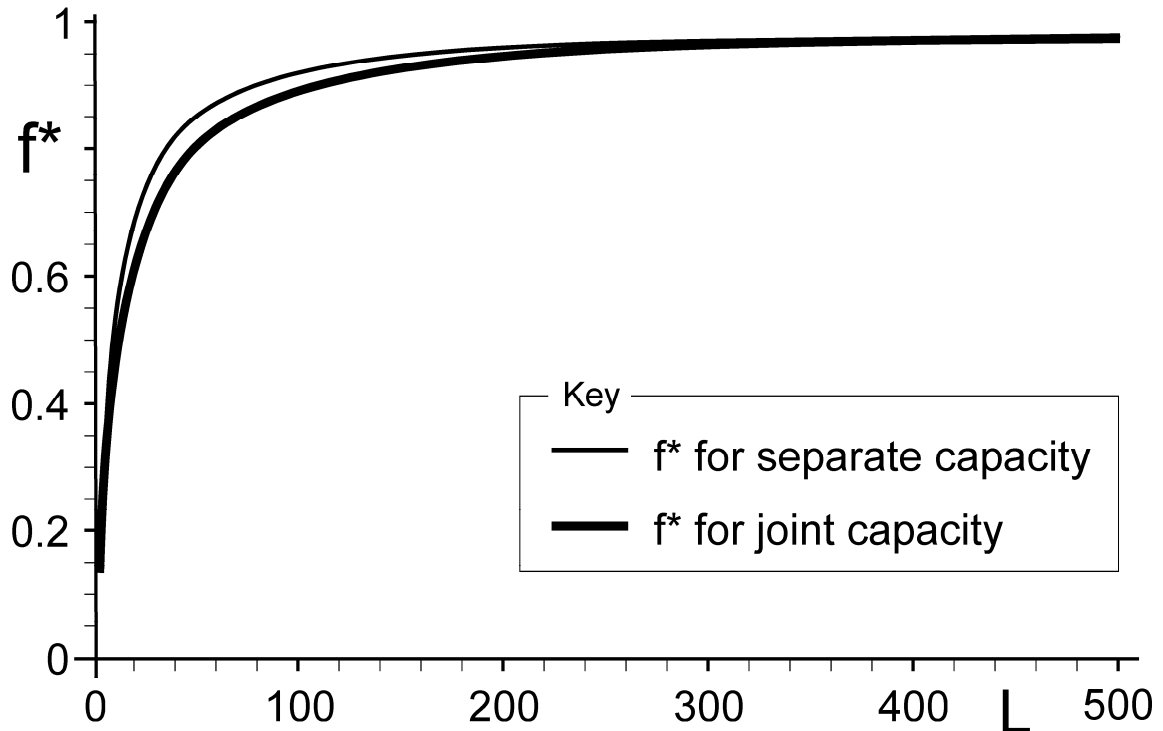


Figure 4. Optimal f for different lead times.

Via a brute force numerical analysis we have some confidence that

$$P^* = L + 1 \quad (25)$$

as every one the of the one million combinations with $L \in \mathbb{N}^+ < 1000$ and $P \in \mathbb{N}^+ < 1000$ resulted in the optimum length of planning cycle conforms to this rule.

3.3. The POUT policy with shared capacity

When the POUT policy is present and the two products share production capacity then the annual costs are given by

$$C = 240 \left(\begin{aligned} & 2\mu U + \frac{w}{P} \sqrt{\frac{fP\sigma^2}{\pi(2-f)}} \exp\left[-\text{erf}^{-1}\left[\frac{w-2U}{w}\right]^2\right] + \\ & (B+H) \sqrt{\frac{2}{\pi}} \sqrt{P\sigma^2 \left(\frac{1}{f(2-f)} + \left\lfloor \frac{L}{P} \right\rfloor\right)} \exp\left[-\text{erf}^{-1}\left[\frac{B-H}{B+H}\right]^2\right] \end{aligned} \right). \quad (26)$$

Taking the derivative w.r.t. f yields

$$\frac{dC}{df} = \frac{240\sigma^2}{(f-2)^2} \left(\begin{aligned} & \frac{W \exp\left[-\text{erf}\left[\frac{w-2U}{w}\right]^2\right]}{\sqrt{\pi} \sqrt{\frac{fP\sigma^2}{f-2}}} + \frac{(f-1)(B+H)P \sqrt{\frac{2}{\pi}} \exp\left[-\text{erf}\left[\frac{B-H}{B+H}\right]^2\right]}{f^2 \sqrt{P\sigma^2 \left(\frac{1}{f(2-f)} + \left\lfloor \frac{L}{P} \right\rfloor\right)}} \end{aligned} \right), \quad (27)$$

which again is a forth order expression in f . We are able to solve for f^* , but the solution is so lengthy that we cannot transpose it in this short conference paper. However we are able to visualise its behaviour for the numerical situation that we used in earlier for the POUT with separate lines. We have also plotted this in

Figure 4. We can see that the f^* for shared capacity is always smaller than the f^* with individual capacity.

Using the same brute force search technique we are able to determine that optimal length of the planning periods with the POUT policy with shared capacity is always one day longer than the real lead-time, L as

$$P^* = L + 1. \quad (28)$$

Finally, we may compare the annual costs for all four scenarios in Figure 5. We can see that the POUT policy with shared production facilities dominates all other situations. The classical OUT policy with individual production lines for each product has the worst economic performance. The POUT policy with separate production lines is more economical than the OUT policy with shared production for small lead-times, but not for lead-times greater than nine.

4. Comparative economic potential of optimising the rescheduling frequency

To see how an optimisation of P compares to other production system improvements, we consider a classical OUT production system, making two products at separate production lines where capacity cannot be shared via shifting labour from one line to the other. We consider the cases where rescheduling initially is done in the daily, weekly, or monthly

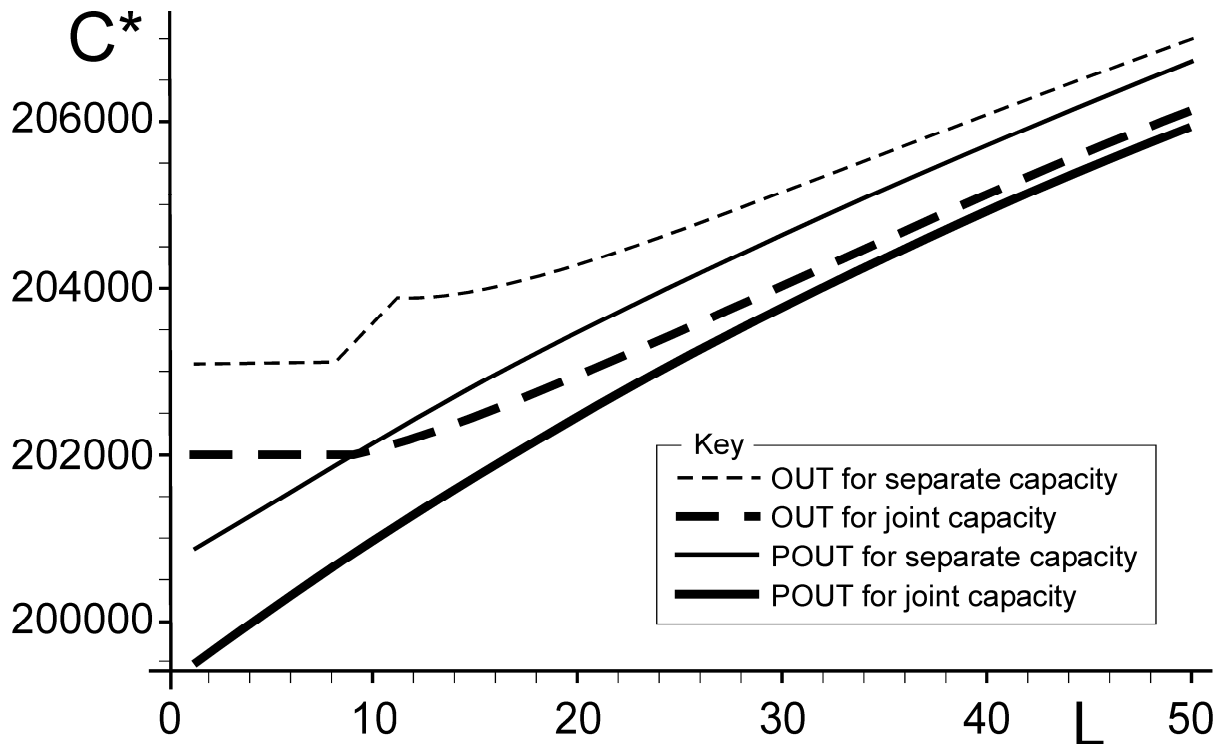


Figure 5. Optimal total cost for different lead times, policies, and production systems.

buckets as these are natural frequencies that fit into the rhythm of life. From these baselines we consider three potential improvements:

1. Consolidation of capacity, so that any unit of labour can produce either product with equal capability. The variability of demand for products *A* and *B* is thus pooled.
2. Application of POUT and optimisation of the f parameter. This seeks to lower total cost by changing the balance between inventory variance and production variance.
3. Changing the rescheduling frequency P to the optimal one, P^* , so that the variance in the production and inventory system can be reduced.

Table 2 shows how the production system responds to all possible combinations of these improvements. It turns out that optimising P gives the single best improvement if we start with weekly or with monthly rescheduling, but starting with daily rescheduling, POUT parameter tuning is more effective. The best system overall uses all of the three improvements. Of particular interest is that the daily system is most sensitive to changes in f , while the system with an optimised P^* is fairly insensitive to this.

The performance of the OUT policy is particularly affected by the choice of P . The POUT policy, on the other hand, sees a smaller economic benefit from optimising P . There is however another strong reason for advocating a P^* configuration for POUT systems, namely the cost function then has a slowly increasing derivative near the minimum, resulting in a more robust system.

5. Discussion and conclusion

Our findings dispel the myth that shorter rescheduling times improve the performance of manufacturing systems, even when no planning costs are present. Instead, there is an optimal rescheduling frequency at which production and inventory control systems should operate. To

Length of planning cycle	Separated Capacity		Consolidated Capacity	
	$f=1$	f^*	$f=1$	f^*
Monthly	207 339	206 196	205 967	205 315
f^*	-	0.628162	-	0.625263
Weekly	209 790	204 929	207 047	203 948
f^*	-	0.203937	-	0.254486
Daily	220 664	204 613	214 530	203 613
f^*	-	0.045914	-	0.059012
Tuned P (21)	204 291	203 410	202 952	202 465
f^*	-	0.628162	-	0.704936

Table 2. Costs for OUT, numerically optimised costs for POUT.

achieve minimum total cost, this period should be either one day longer than the physical lead time, or at a point that is defined by the costs involved as given by (18).

Our work has been involved with systems that have large production volumes; we have not investigated the specific case of low-volume production. Set-up / changeover costs are also ignored, these may otherwise be an important, if not dominating factor.

We have demonstrated that using a single production line for several products is preferable to running each product on a separate line, as this further exploits the demand pooling effect that is present. In our example, the savings made possible by consolidating production (or using multi-skilled labour) in this way are less than those of optimising the rescheduling frequency, but are still significant. As a final improvement to the system, the adoption of the POUT policy gives a further reduction of cost, albeit smaller than from the other efforts.

Importantly there may be severe economic penalties involved when the planning cycle is too short and that these systems have an enhanced sensitivity to errors in setting an appropriate value for the proportional feedback controller in the POUT policy. The OUT policy, with a unity feedback controller, suffers particularly, as costs increase quite significantly when the rescheduling period is reduced beyond its optimal length.

Interestingly the costs for an OUT system are never lower than the costs for an equivalent and optimised POUT system, while optimal POUT systems never reschedule more frequently than their OUT counterparts. This suggests that quickly responding systems are not necessarily more cost-efficient than slower ones. The fastest systems are plagued by higher costs and cannot easily compete with systems that have optimal rescheduling.

As the optimum rescheduling cycle is always at least one day longer than the lead time, it follows that T_p should be zero, meaning that these policies do not need to consider work-in-process (WIP) inventory when order rates are calculated. This obviates the need to monitor WIP (or goods in transit – which is often a rather difficult thing to do in an industrial setting) and at the same time allows the production and inventory system to perform optimally.

6. References

Bicheno, J. and Holweg, M. 2009. *The Lean Toolbox 4th Edition: The Essential Guide to Lean Transformation*. PICSIE Books, Buckingham, UK.

- Burbidge, J.L. 1988. *Production Flow Analysis for Planning Group Technology*. Oxford University Press, Oxford, UK.
- Burbidge, J.L. 1983. Five golden rules to avoid bankruptcy. *Production Engineer*. 62 (10). 13-14.
- Disney, S.M., Gaalman, G. and Hosoda, T. 2012. "Review of stochastic cost functions for production and inventory control", *Pre-prints of the 17th International Working Seminar of Production Economics*, Innsbruck, Austria, February 20th -24th.
- Disney, S.M., Towill, D.R., van de Velde, W., 2004. Variance amplification and the golden ratio in production and inventory control. *International Journal of Production Economics*, 90, 295-309.
- 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.
- Maister, D.H. 1976. "Centralization of Inventories and the 'Square Root Law'" *International Journal of Physical Distribution*, 6(3), 124-134.
- Ohno, T., 1988. *Toyota Production System: Beyond Large-scale Manufacturing*. Productivity Press, New York, USA.
- Shingo, S., 1989. *The Toyota Production System from an Industrial Engineering Viewpoint*. Productivity Press, New York, USA.
- Sloan, A.P., 1963. *My Years with General Motors*. Doubleday, New York, USA.