

Mitigating Variance Amplification Under Stochastic Lead-Time: The Proportional Control Approach

Xun Wang^a, Stephen M. Disney^{a,*}

^a*Logistics Systems Dynamics Group, Cardiff Business School, Cardiff University, UK.*

Abstract

Logistic volatility is a significant contributor to supply chain inefficiency. In this paper we investigate the amplification of order and inventory fluctuations in a state-space supply chain model with stochastic lead-time, general auto-correlated demand and a proportional order-up-to replenishment policy. We identify the exact distribution functions of the orders and the inventory levels. We give conditions for the ability of proportional control mechanism to simultaneously reduce inventory and order variances. For AR(2) and ARMA(1,1) demand, we show that both variances can be lowered together under the proportional order-up-to policy. Simulation with real demand and lead-time data also confirms a cost benefit exists.

Keywords: Inventory, bullwhip effect, stochastic lead-time, demand correlation

1. Introduction

We investigate the performance of the order-up-to (OUT) and proportional order-up-to (POUT) inventory control policies via the variance of the inventory and orders under a stochastic lead-time. Variability in inventory systems is commonly generated by uncertainties in demand, supply, transportation, and manufacturing. This variability can be amplified by poorly designed replenishment policies (Lee et al., 1997; Chen et al., 2000). Fluctuations in the replenishment orders and inventory levels pose an operational threat to companies. High order variance (a.k.a. the bullwhip effect) brings uncertainty to the upstream supplier,

*Corresponding author.

Email addresses: WangX46@cardiff.ac.uk (Xun Wang), DisneySM@cardiff.ac.uk (Stephen M. Disney)

and reduces supply chain efficiency. Similarly, high inventory variance results in high safety stock levels and/or poor customer service, which in turn leads to inflated inventory cost.

Logistics uncertainty and stochastic shipping delays are a major component of supply chain risk. In recent years, production and distribution systems have become increasingly global, exposing supply chains to more volatility than ever before. Global transportation routes, involving air, truck, rail and ocean freight modes, have long and variable lead-times, due to external factors such as seasonality effects, security and customs delays and slow steaming. Uncertain lead-times sometimes trigger another effect called *order crossover*, where replenishments are received in a different sequence than they were ordered. Whilst these two concepts do not necessarily imply each other (Zipkin, 1986; Riezebos, 2006), a highly variable lead-time often results in order crossover. This is especially so in global supply chains where container liners may take different routes, overtake each other at sea, and stop at different ports along the way. Furthermore, individual containers may be held up for customs inspections at national borders.

Hayya et al. (2008a) classified the research on stochastic lead-time into three schools: the Hadley-Whitin School (Hadley and Whitin, 1963), which assumes that the probability of order crossover is so small that it can be totally ignored; the Zipkin-Song School (Zipkin, 1986; Song, 1994), which assumes that goods are processed sequentially (perhaps in some sort of first-in-first-out queue) so that order crossover cannot happen; and the Zalkind School (Zalkind, 1978; Robinson et al., 2001; Bradley and Robinson, 2005), which takes order crossover into account and discovers that inventory cost and safety stock can be reduced by considering this effect. Models of this kind are first introduced by Finch (1961) and Agin (1966), which gave the correct expression for the distribution of the number of outstanding orders. Zalkind (1978) determined the optimal target inventory level to minimize total cost. Bagchi et al. (1986) showed the importance of considering order crossover when setting safety stock. Robinson et al. (2001) highlighted that order crossover has a significant impact on inventory control and should not be ignored. The aims of these studies are either to derive (approximate) relevant distributions or to decide safety stock parameters.

In recent years the impact of order cross-over on inventory management is gaining aca-

demic attention. Chatfield et al. (2004) and Kim et al. (2006) have investigated the bullwhip effect with stochastic lead-time, adopting the assumptions of i.i.d. demand and the OUT replenishment policy. Hayya et al. (2008b) considered the inventory cost optimization problem under order crossover using regression on empirical data. Hayya et al. (2011) further studied the impact of order-crossover on inventory cost, assuming deterministic demand and exponentially distributed lead-time. Bischak et al. (2014) showed that taking into account order crossover and using an approximate *effective lead-time* deviation allows companies to reduce inventory costs.

Another stream of research has shown that the POUT policy is effective at smoothing the bullwhip effect at a cost of increased inventory variability (Gaalman, 2006; Chen and Disney, 2007). However, most studies on bullwhip effect require at least a predictable, if not a constant, lead-time; while existing research on stochastic lead-time problems do not explicitly tackle the amplification problem. Disney et al. (2016) tried to fill this gap by considering an inventory system with stochastic lead-time and order-crossover. They derived the distribution of orders and inventory under stochastic lead-time and discussed the impact of the proportional OUT policy on costs and safety stocks. However, the demand pattern is restricted to i.i.d. and no formal proof is given for the superiority of the POUT policy over the OUT policy.

This paper is a sequel to Disney et al. (2016) in which we extend, sharpen and refine their results in the following ways: (1) we identify the distributions of order and inventory under stochastic lead-time and auto-correlated demand; (2) we provide conditions when the OUT and POUT policies minimize inventory variability under the ARMA(p,q) demand process; (3) we examine the possibility of simultaneous reduction of inventory and order variances by proportional control. Below we list our contributions in more detail.

- We develop a state space approach which allows us to derive the probability density functions of orders and inventory under the POUT policy, arbitrarily distributed stochastic lead-time and general ARMA(p,q) demand. The pdfs then allow us to derive exact expressions for the inventory and order variances.

- We give a necessary condition for when the OUT policy minimizes the inventory variance under general ARMA demand and a stochastic lead-time. Based on this condition we prove that the OUT policy is never optimal for minimizing inventory variance when order crossover is present and demand is temporally independent.
- We give a precise condition under which the inventory and order variances can be reduced simultaneously by optimizing the proportional controller in the POUT policy. Parametrical combinations for this condition are derived for special cases of AR(2) and ARMA(1,1) demand. Simultaneous reduction of inventory and order variance via proportional control is possible for the majority of demand processes.

The paper is organized as follows. In §2 we introduce notation and modelling basics. §3 contains the main results, which includes an exact approach to obtain the distribution of order and inventory, conditions for the optimality of the OUT policy, and conditions for the simultaneous improvement of inventory and order variances. In §4 we numerically investigate the impact of demand correlation and lead-time uncertainty. The cost implications of the proportional policy are also provided based on real demand and lead-time data. Finally we conclude and discuss our results in §5. Proofs that are not outlined in the main text are presented in the Appendix.

2. Modelling the demand and ordering policy

In this section we establish the model including the objective function, demand process, forecasting and inventory control policies, sequence of events, and the balance equations. We focus on a periodic review inventory system where the system states are defined on \mathbb{R} . The lead-time, defined on \mathbb{N}^+ , is a positive random variable following any arbitrary non-negative discrete distribution that is independent over time. The assumption of discrete lead-time is natural as the lead-time is measured in units of the review period in periodic systems (Disney et al., 2016). Since our model allows for order crossovers, there are no restrictions on the lead-times of consecutive orders. The demand is a normally distributed ARMA(p,q)

process. Both the lead-time distribution and demand correlation are known in advance. In practice this knowledge can be realized by statistically analysing historical data.

Table 1 lists commonly used notation. Importantly we denote $\Sigma_{xy}(\tau)$ as the mutual covariance function between the random variables x and y , with time difference τ . $E(x)$ or μ_x is the expectation of x . Variables x and y don't have to be scalars. If x and y each contains m and n scalar random variables respectively, $\Sigma_{xy}(\tau)$ is an $m \times n$ matrix and $E(x)$ is a $1 \times m$ vector. When $\tau = 0$, $\Sigma_{xx}(0)$ is the autocovariance matrix of x . Sometimes we write this as Σ_{xx} if no other confusion would occur. The leading diagonal of Σ_{xx} contains the variances of the elements in x . Other notation will be introduced when necessary.

2.1. Sequence of events and the balance equations

We consider a 3 node supply chain model consisting of an end consumer, a manufacturer and a supplier. The consumers demand is assumed to be exogenous, and the supplier is never capacitated. We define the sequence of events as follows. In each period t , the supplier ships the order it has received from the manufacturer in the previous period, and the manufacturer receives the shipment corresponding to the order placed $t - L$ periods ago, where L is the (possibly stochastic) lead-time. Next, the customer demand is observed and satisfied. The WIP and inventory information is updated. In our definition, i_t refers to the *net* inventory level after order completion and demand consumption in period t , hence $i_t < 0$ denotes a backlogged situation. Finally, the manufacturer calculates a forecast and issues a new order based on updated forecast, inventory and WIP information.

The balance equations for inventory, i_t , and WIP, w_t , in our model are

$$\begin{aligned} i_t &= i_{t-1} + o_{t-L} - d_t, \\ w_t &= w_{t-1} + o_{t-1} - o_{t-L}. \end{aligned} \tag{1}$$

Here o_t is the replenishment order placed at time t . Under a stochastic lead-time there could be zero or multiple orders arriving in one period. For modelling purposes, we assign a lead-time value to an order at the time when the order is issued. This is to ensure that an order is only received once. Here we also assume that an order will always be received in full, regardless of the lead-time.

Table 1: Commonly used notation in this paper

Variables (time-dependent random processes)	
y_t, z_t	Variables in the ARMA model
ε_t	Gaussian i.i.d. variable with zero mean and unit variance
d_t	Demand
\hat{d}_t	Demand forecast
\hat{D}_t	Lead-time demand forecast
S_t	Order-up-to level
o_t	Order
i_t	Net inventory level
w_t	Work-in-process
IP_t	Inventory position, $IP_t = i_t + w_t$
ξ_t	Pipeline Status
$x_t(\xi)$	Sub-process of $\{x_t\}$ with pipeline status ξ
Parameters (constant)	
ϕ, θ	Auto-correlation and moving average parameters in the ARMA demand model
ss	Safety stock level
$1 - \lambda$	Proportional feedback controller
L^+, L^-	Maximum and minimum lead-time
h, b	Unit holding and backlog cost
First-order moment	
$E(x), \mu_x$	Expectation of x
$E(x; \xi)$	Expectation of $x_t(\xi)$
Second-order moments	
$\Sigma_{xy}(\tau)$	Mutual covariance function between x and y with time difference τ
$\Sigma_{xx}(0), \Sigma_{xx}$	Autocovariance matrix of x
$\Sigma_{xy}(0; \xi)$	Mutual covariance between $x_t(\xi)$ and $y_t(\xi)$
Probabilities and distribution functions	
p_L	Probability of lead-time being L periods long
$\psi_x(\cdot)$	Probability density function of x
$\Psi_x(\cdot)$	Cumulative distribution function of x
$\bar{\Psi}_x(\cdot)$	Complementary cumulative distribution function of x
$\varphi(x \mu, \sigma^2)$	Probability density function of normal distributed variable x with mean μ and variance σ^2
Matrices	
I	Appropriately dimensioned identity matrix
$\mathbf{1}$	Appropriately dimensioned unit column vector
A^T	Transpose of A
A^{-1}	Inverse of A
$\text{diag}\{\cdot \cdot \cdot\}$	Block-diagonal matrix

2.2. The objective functions

We consider the minimization of the inventory variance, Σ_{ii} , as the primary objective. The reason is threefold. First, from a practical perspective, a smooth inventory process is much easier to manage than a highly fluctuating one. An inventory process with low variance requires less safety stock, which leads to low inventory and backlog levels. Second, from a modelling perspective, under mild assumptions, minimizing the variance of the inventory is equivalent to minimizing the inventory cost. If the lead-time is constant, then the optimized piecewise-linear inventory cost is proportional to the standard deviation of the inventory levels. Third, under the quadratic assumption, the problem defined in this paper becomes a linear quadratic Gaussian (LQG) problem which greatly promotes the tractability of the problem.

The variance of orders, Σ_{oo} , serves as a second objective due to the large negative impact on supply chain inefficiency brought about by volatile production rates and replenishment orders. By means of the proportional policy, we do not try to minimize Σ_{oo} due to its triviality (setting all orders to the mean demand ensures $\Sigma_{oo} = 0$). Instead, we will evaluate Σ_{oo} under an optimal POUT policy for minimizing Σ_{ii} .

In §4 we numerically investigate the inventory cost under the assumption of a piecewise linear cost function. The expected cost C is defined as the expected inventory holding and backlog costs, which are proportional to the inventory level and the backlog level respectively,

$$C = hE(i_t^+) + bE(-i_t)^+, \quad (2)$$

where $x^+ = \max(x, 0)$, i_t^+ is the inventory level in period t when $i_t > 0$; $(-i_t)^+$ is the backlog level when $i_t < 0$.

2.3. The demand process

An ARMA(p, q) demand, $d_t = \mu_d + z_t$, can be formulated as a constant, μ_d , plus a zero mean ARMA(p, q) process, z_t , where

$$z_t = \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}.$$

Here $\{\varepsilon_t\}$ is a Gaussian white noise process. Let $m = \max(p, q + 1)$ and $\theta_j = 0$ for $j = q + 1, \dots, m$; $\phi_j = 0$ for $j = p + 1, \dots, m$. Introduce another $1 \times m$ vector, y , such that

$$y_{j,t} = \phi_j z_{t-1} + \dots + \phi_m z_{t+j-m-1} - \theta_j \varepsilon_{t-1} - \dots - \theta_m \varepsilon_{t+j-m-1},$$

where $y_{j,t}$ is the j th element of y_t , $j = 1, 2, \dots, m$. With this definition we can transform the ARMA demand model into a canonical state space form

$$\begin{aligned} y_t &= Ay_{t-1} + B\varepsilon_t, \\ z_t &= Cy_t. \end{aligned} \tag{3}$$

Here A is an $m \times m$ left companion matrix with $A_{j1} = \phi_j$ and $A_{j,j+1} = 1$. $B = (1 \ -\theta_1 \ \dots \ -\theta_{m-1})^T$ and $C = (1 \ 0 \ \dots \ 0)$.

The auto correlation of y satisfies

$$\Sigma_{yy} = A\Sigma_{yy}A^T + BB^T. \tag{4}$$

The solution of (4) can be found from

$$\text{vec}(\Sigma_{yy}) = (I - A \otimes A)^{-1} \text{vec}(BB^T), \tag{5}$$

in which \otimes denotes the Kronecker (tensor) product of matrices, $\text{vec}(\cdot)$ is the matrix column-wise stacking operation which transforms an $m \times n$ matrix into an $(m \times n) \times 1$ column vector, such that for matrix $A = (a_1 \ a_2 \ \dots \ a_n)$ where a_1, a_2, \dots, a_n are $m \times 1$ column vectors, $\text{vec}(A) = (a_1^T \ a_2^T \ \dots \ a_n^T)^T$.

2.4. The forecasting and inventory control policies

When the lead-time is longer than the review period, the manufacturer has to determine the forecast of the lead-time demand \hat{D} . This is a sum of all k -step-ahead demand forecasts over the lead-time, $\hat{D}_t = \sum_{k=1}^L \hat{d}_{t,k}$. $\hat{d}_{t,k}$ is the conditional expectation of d_{t+k} based on all observations available at time t which the manufacturer uses to produce a minimum mean squared error (MMSE) k -step-ahead forecast of demand. It was chosen as minimizing the squared forecast errors over the lead-time and review period results in minimal inventory

variance (Hosoda and Disney, 2006). Although MMSE forecasting requires the full knowledge of the demand process, we note that our approach can be easily extended to incorporate any linear forecasting method.

To calculate $\hat{d}_{t,k}$ we need to first derive the corresponding forecast of ARMA state variable y . The MMSE k -step-ahead forecast for y is given by

$$\hat{y}_{t,k} = A^k y_t, \quad (6)$$

where A^k is the matrix A raised to its k th power. The k -step ahead forecasts of d are then simply $\hat{d}_{t,k} = C\hat{y}_{t,k} + \mu_d$.

We study the POUT policy, which can be formed by adding a proportional controller into the inventory position feedback loop of the OUT policy. This policy has been studied by Magee (1956); Deziel and Eilon (1967); John et al. (1994); Disney et al. (2004). We start from the classic OUT policy under constant lead-time before introducing the POUT policy under stochastic lead-time. The OUT policy is defined as

$$o_t = S_t - IP_t;$$

$$S_t = \hat{D}_t + ss.$$

This policy can be further rewritten as (Dejonckheere et al., 2003)

$$o_t = \hat{d}_{t,L} + (ss + \hat{w}_t - i_t - w_t), \quad (7)$$

where \hat{w}_t is the (time varying) target WIP level, $\hat{w}_t = \sum_{k=1}^{L-1} \hat{d}_{t,k}$. w_t is the actual WIP level and under a constant lead-time, $w_t = \sum_{k=1}^{L-1} o_{t-k}$. ss is the target inventory level.

The POUT policy is formed by adding a proportional controller, $1 - \lambda$, to the second term of (7) such that

$$o_t = \hat{d}_{t,L} + (1 - \lambda)(ss + \hat{w}_t - i_t - w_t). \quad (8)$$

The proportional controller must satisfy $\lambda \in (-1, 1)$ for stability (Disney, 2008). Note that $ss + \hat{w}_t$ is the target *inventory position* and $i_t + w_t$ is the actual inventory position. In other words, the order quantity equals the L -step-ahead demand forecast plus a fraction $(1 - \lambda)$

of the discrepancy between the target and actual inventory position. Since the lead-time is i.i.d. with a known discrete distribution, $\hat{d}_{t,L}$ and \hat{w}_t can then be calculated as the average over all possible lead-time values, which we denote as:

$$\begin{aligned}\hat{d}_{t,L} &= \sum_{k=L^-}^{L^+} p_k \hat{d}_{t,k}, \\ \hat{w}_t &= \sum_{l=L^-}^{L^+} p_l \sum_{k=1}^{l-1} \hat{d}_{t,k}.\end{aligned}\tag{9}$$

Rearranging (8) yields

$$o_t = f_t + (1 - \lambda)(ss - i_t - w_t),\tag{10}$$

where f_t is the forecast term

$$f_t = \hat{d}_{t,L} + (1 - \lambda)\hat{w}_t = \sum_{k=L^-}^{L^+} p_k \hat{d}_{t,k} + (1 - \lambda) \sum_{l=L^-}^{L^+} p_l \sum_{k=1}^{l-1} \hat{d}_{t,k}.\tag{11}$$

From (6) we see that the MMSE forecast is linear, which enables us to rewrite f_t as a linear function of y_t , i.e., $f_t = Fy_t + (\mu_L - \lambda\mu_L + \lambda)\mu_d$. F is the forecasting vector,

$$F = \sum_{k=L^-}^{L^+} p_k CA^k + (1 - \lambda) \sum_{l=L^-}^{L^+} p_l \sum_{k=1}^{l-1} CA^k.$$

The variance of f takes the quadratic form

$$\Sigma_{ff} = F\Sigma_{yy}F^T.$$

3. Distributions of the orders and the inventory

3.1. Revisiting the constant lead-time case

To reveal the impact of stochastic lead-time, we first revisit some well-established results under constant lead-time to highlight the trade-off between order and inventory variance. The demand is assumed to be i.i.d. (white noise) here, but the trade-off also exists under auto-correlated demand.

Lemma 1. *Under i.i.d. demand, a constant lead-time L and the POUT policy (8), the order and inventory variances are*

$$\Sigma_{oo} = \frac{1 - \lambda}{1 + \lambda} \Sigma_{dd},$$

and

$$\Sigma_{ii} = \left(\frac{\lambda^2}{1 - \lambda^2} + L \right) \Sigma_{dd}.$$

Proof. A proof was provided in Disney et al. (2004). □

The following observations can be made from Lemma 1. (1) *Order-up-to optimality:* the optimal λ for minimizing Σ_{ii} is $\lambda = 0$. In other words, the OOT policy minimizes Σ_{ii} . (2) *Variance trade-off:* Under a constant lead-time, at $\lambda = 0$, Σ_{oo} can be reduced only at a cost of increasing Σ_{ii} .

3.2. Distribution of orders under a stochastic lead-time

In this section we characterise the order distribution, including the type of distribution, the mean and the (co)variance functions under a stochastic lead-time.

Lemma 2. *For the inventory evolution given by (1), demand processes by (3) and the ordering policy by (8), under a stochastic lead-time, the orders and the inventory position is normally distributed with $E(o) = \mu_d$ and $E(IP) = ss + (\mu_L - 1)\mu_d$.*

The following Lemma gives the auto- and mutual- covariance function between y and o .

Lemma 3. *For the inventory evolution given by (1), demand processes by (3) and the ordering policy by (8), under a stochastic lead-time we have:*

Covariance between the demand and orders,

$$\Sigma_{yo} = (I - \lambda A)^{-1} [(I - A)\Sigma_{yy}F^T + (1 - \lambda)\Sigma_{yy}C^T]. \quad (12)$$

Mutual covariance function between demand and orders,

$$\Sigma_{yo}(\tau) = A^\tau \Sigma_{yo}(0). \quad (13)$$

Variance of orders,

$$\begin{aligned} \Sigma_{oo} = & \frac{2}{1+\lambda} F(I - \lambda A)^{-1} (I - A) \Sigma_{yy} F^T + \frac{2\lambda}{1+\lambda} F(A - I)(I - \lambda A)^{-1} \Sigma_{yy} C^T + \\ & \frac{2}{1+\lambda} C(I - \lambda A)^{-1} (I - A) \Sigma_{yy} F^T + \frac{1-\lambda}{1+\lambda} C(I + \lambda A)(I - \lambda A)^{-1} \Sigma_{yy} C^T. \end{aligned} \quad (14)$$

Autocovariance function of orders,

$$\Sigma_{oo}(\tau) = [FA + (1 - \lambda)CA - F] \Sigma_{yo}(\tau - 1) + \lambda \Sigma_{oo}(\tau - 1). \quad (15)$$

The probability density function of orders $\psi_o(x)$ can be characterized as,

$$\psi_o(x) = \varphi(x | \mu_d, \Sigma_{oo}).$$

From Lemma 2 we see that $E(o)$ is not affected by the lead-time; but Lemma 3 shows Σ_{oo} is affected by the lead-time distribution through the forecasting vector F . More importantly, the process of the orders is stationary. Lemma 3 not only gives the variance of orders, but also provides an iterative approach to calculate the auto/mutual covariance functions of y and o which is essential in calculating the inventory distribution.

3.3. Distribution of inventory under a stochastic lead-time

In a constant lead-time scenario, whether or not a previous order o_{t-k} is outstanding at time t is deterministic. However, if lead-time is stochastic then this status becomes probabilistic. To derive the distribution of inventory we have to understand how it is affected by the stochastic lead-time in terms of the outstanding order status. Two definitions are necessary here.

Definition 1. The *pipeline status* ξ is defined as a $1 \times (L^+ - 1)$ binary vector. The k th element of ξ , $\xi(k)$, is a random binary variable for $k \geq L^-$ and $\xi(k) = 1$ for $k < L^-$.

ξ_t indicates which orders are outstanding and which have been received at time t . Specifically $\xi_t(k) = 1$ means the order placed at time $t - k$ is outstanding at t and $\xi_t(k) = 0$ means the order placed at $t - k$ is completed (or has been received) at t , $k = 1, 2, \dots, L^+ - 1$. Note that for any t , orders made before (and including) $t - L^+$ are fulfilled, i.e., $\xi_t(k) \equiv 0$ for

Table 2: Example of stochastic lead time and outstanding orders

Period	1	2	3	4	5	6	7	8	9	10
Order lead-time	3	3	1	3	2	1	1	3	1	2
Outstanding orders	-	-	o_1, o_2	o_2	o_4	o_4, o_5	None	None	o_8	o_8
Pipeline status	-	-	(1, 1)	(1, 0)	(0, 1)	(1, 1)	(0, 0)	(0, 0)	(0, 1)	(1, 0)

$k \geq L^+$, therefore they are not included in the pipeline. Orders made after (and excluding) $t - L^-$ are outstanding, i.e., $\xi_t(k) \equiv 1$ for $k < L^-$. So the pipeline contains a maximum of $L^+ - 1$ outstanding orders, in which the first $L^- - 1$ orders have deterministic status (they are all outstanding) and the other $L^+ - L^-$ orders have random status. ξ has $2^{L^+ - L^-}$ possible realizations, which we index as ξ^j . Here the superscript $j = 1, 2, \dots, 2^{L^+ - L^-}$ denotes all possible realizations of vector ξ , which should not be confused with $\xi(k)$, the k^{th} element of ξ .

Definition 2. A *sub-process* $\{x_t(\xi)\}$ of a process $\{x_t\}$ is defined as a subset of the process when the pipeline status is unique, e.g., $\{i_t(\xi^j)\} = \{i_t | \xi_t = \xi^j\}$, $\{w_t(\xi^j)\} = \{w_t | \xi_t = \xi^j\}$.

Example. Consider the lead-time distribution: $p_k = 1/3$, $k = 1, 2, 3$. The second line in Table 2 is a realization of the stochastic lead-time for each order in the 10 period simulation. As $L^+ = 3$ we start our investigation from period 3. Since the lead-time for o_1 and o_2 are both 3, in period 3 they are both outstanding. In period 4, o_1 has arrived and so does o_3 , which has a lead-time of 1. The only outstanding order is o_2 . At period 5, o_2 has arrived and o_4 is outstanding. This constitutes the third line of Table 2. The WIP with the same pipeline status belongs to one sub-process, e.g., periods 3 and 6, periods 4 and 10, period 5 and 9, periods 7 and 8. Thus there are several sub-processes (4 in this example as $2^{3-1} = 4$) in the WIP process. Each of the sub-processes is characterized by a pipeline status vector, (1, 1), (1, 0), (0, 1) and (0, 0), shown in the last row of Table 2.

Equation (10) reveals that the distribution of inventory is a convolution of the distribution functions of f_t , o_t and w_t , and both f_t and o_t are independent of the lead-time realization. Hence we can decompose the inventory process into several sub-processes as-

sociated with the different pipeline status. By such division, the random variables in one sub-process follow identical normal distributions, due to the fixed pipeline status. The complication of the current problem, compared with the assumptions of i.i.d. demand and OUT policy in the literature (Robinson et al., 2001), is that the order process is now correlated as a consequence of the demand correlation and the proportional feedback controller, so the covariances must be accounted for. However, the covariance functions with the same pipeline status will be identical.

For the inventory sub-process $\{i_t(\xi)\}$, we have the following Lemma:

Lemma 4. *The mean and variance of $\{i_t(\xi)\}$ are given by*

$$E(i; \xi) = ss + (\mu_L - \xi \mathbf{1} - 1)\mu_d,$$

and

$$\Sigma_{ii}(0; \xi) = \frac{\Sigma_{ff} + \Sigma_{oo} - 2F\Sigma_{yo}}{(1 - \lambda)^2} + \frac{2\xi(\alpha - \beta)}{1 - \lambda} + \xi\Gamma\xi^T.$$

where α and β are $(L^+ - 1) \times 1$ column vectors, Γ is a $(L^+ - 1) \times (L^+ - 1)$ covariance matrix of orders, in which the elements are $\alpha_j = \Sigma_{oo}(j)$, $\beta_j = F\Sigma_{yo}(j)$, $\gamma_{jk} = \Sigma_{oo}(j - k)$ respectively, $j, k = 1, 2, \dots, L^+ - 1$.

We use $p(\xi)$ to denote the probability of ξ , the proportion of the sub-process characterized by ξ in the whole process. Zalkind (1978) shows that

$$p(\xi) = \prod_{k=1}^{L^+-1} \{[1 - \xi(k)] \Psi_L(k) + \xi(k) \bar{\Psi}_L(k)\}. \quad (16)$$

where $\Psi_L(\cdot)$ and $\bar{\Psi}_L(\cdot)$ are the cumulative distribution function and the complementary cumulative distribution function of the discrete lead-time. The full inventory process is a simple mixture of the sub-processes, which are mutually independent (because the lead-time is temporally independent). Therefore after determining the first- and second- order moments of the sub-processes, the distribution function of the full inventory process can be derived by summing over the distributions of all the sub-processes:

$$\psi_i(x) = \sum_{j=1}^{2^{L^+ - L^-}} p(\xi^j) \varphi(x | E(i; \xi^j), \Sigma_{ii}(0; \xi^j)). \quad (17)$$

We can now obtain a general expression for the inventory variance.

Proposition 1. *The variance of $\{i_t\}$ is given by*

$$\Sigma_{ii} = \frac{\Sigma_{ff} + \Sigma_{oo} - 2F\Sigma_{yo}}{(1-\lambda)^2} + \sum_{k=1}^{L^+-1} \bar{\Psi}_L(k) \left[\Sigma_{oo} + 2 \frac{\Sigma_{oo}(k) - F\Sigma_{yo}(k)}{1-\lambda} \right] + \sum_{k=1}^{L^+-2} \nu_k \Sigma_{oo}(k) + \mu_d^2 \Sigma_{NN}, \quad (18)$$

where

$$\nu_k := 2 \sum_{j=1}^{L^+-k-1} \bar{\Psi}_L(j) \bar{\Psi}_L(j+k),$$

and the variance of the number of outstanding orders

$$\Sigma_{NN} = \sum_{j>k} p(\xi^j) p(\xi^k) (\xi^j \mathbf{1} - \xi^k \mathbf{1})^2.$$

The first term of (18) is the inventory variance when $L = 1$. The last term gives the inventory variance under constant demand, i.e., $d_t \equiv \mu_d$. It can be seen that a larger mean demand increases the inventory variance under stochastic lead-times. This is contrary to the deterministic lead-time case where the mean demand does not affect the inventory variance, but the variance of inventory is greatly affected by the mean demand under a stochastic lead-time. Moreover, since $\psi_i(x)$ is a sum of several normal distributions with different means, it does not have to be normally distributed or even unimodal.

Also note that (18) is different from the inventory variance expression in Disney et al. (2016) in the following ways. First, the demand auto-correlation is taken into account, which affects all the covariance terms, especially those related to the forecast. Second, the variance expression is based on the lead-time distribution directly, rather than on the probabilities of ξ . The latter improvement is essential in proving the sub-optimality of the OUT policy in §3.4.

Our approach to calculate inventory variance needs the lead-time to be bounded, which is practically plausible. If the lead-time follows a theoretical distribution where lead-time is unbounded, computational difficulties may arise. As noted by Robinson et al. (2001) this would require calculating an infinite sum of random variables. However, under the assumption of $\lim_{L \rightarrow \infty} p_L = 0$, we anticipate that our approach can generate an approximation to any desired level of accuracy.

3.4. Conditions for optimality and variance trade-off

Having extended the lead-time assumption from deterministic to stochastic, the following questions naturally arises: (1) will the OUT policy still be the optimal policy to minimize inventory variance? (2) If the answer to (1) is no, then will there still be variance trade-off in the OUT policy, i.e., what happens to the order variance as we optimize inventory variance? The non-optimality of order-up-to policy means that there is at least a $\lambda \neq 0$ which gives lower inventory variance; the non-existence of the variance trade-off means that the order variance can be reduced simultaneously with inventory variance – a highly desirable situation. Hereafter we denote λ^* as the optimal feedback controller for minimizing the inventory variance. Then these two questions can be reiterated as: (1) Is $\lambda^* = 0$? (2) Does $\Sigma_{oo}|_{\lambda=\lambda^*} > \Sigma_{oo}|_{\lambda=0}$?

Based on the explicit expression of inventory variance, we now present several results that answer these questions. First, we give a necessary condition for the OUT policy to be optimal at minimizing the inventory variance. Moreover, we give exact conditions under which Σ_{ii} and Σ_{oo} can be reduced together. For the special case of i.i.d. demand, we can derive a stronger result, that is, if there is order cross-over present, i.e., the difference between maximum and minimum i.i.d. lead-time is larger than 1, then the OUT policy is never optimal in minimizing inventory variance; and the optimal proportional policy will always lead to a reduced order variance. This phenomenon is observed in Disney et al. (2016) and we prove it here.

Lemma 5. *The derivative of the order variance w.r.t. λ is*

$$\frac{\partial \Sigma_{oo}}{\partial \lambda} = 2QP_0\Delta Q^T, \quad (19)$$

and the derivative of the inventory variance w.r.t. λ is

$$\frac{\partial \Sigma_{ii}}{\partial \lambda} = Q \left(P_1 + 2 \sum_{k=1}^{L^+-1} \bar{\Psi}_L(k) P_{2k} + \sum_{k=1}^{L^+-2} \nu_k P_{3k} \right) \Delta Q^T, \quad (20)$$

where $Q := (F \quad \partial F / \partial \lambda \quad C)$, $\Delta := \text{diag}\{\Sigma_{yy}, \Sigma_{yy}, \Sigma_{yy}\}$,

$$P_0 := \begin{pmatrix} -(I-A)^2 & I-A & A-I \\ I-A & 0 & 0 \\ -(I-A)^2 & I-A & A-I \end{pmatrix},$$

$$P_1 := \begin{pmatrix} 2A & I & -4I \\ I & 0 & -2I \\ 2(I-A^2) & 2(I-A) & 2A \end{pmatrix},$$

$$P_2 := \begin{cases} \begin{pmatrix} 0 & 0 & -I \\ 0 & 0 & -I \\ I-A^3 & I-A^2 & A^2 \end{pmatrix} & k=1 \\ \begin{pmatrix} (A^k + A^{k-2} - A + I)(A-I) & (A^{k-1} - I)(A-I) & -A^k + A^{k-1} - A^{k-2} + A - I \\ (A^{k-1} - I)(A-I) & 0 & -A^{k-1} \\ (A^{k+1} + A^{k-1} + A - I)(I-A) & (A^{k-1} + I)(I-A) & A^{k+1} - A^k + A^{k-1} + A - I \end{pmatrix} & k > 1 \end{cases}$$

and

$$P_3 := \begin{cases} \begin{pmatrix} (A^2 - A + 2I)(I-A) & -(I-A)^2 & (I-A)^2 \\ -(I-A)^2 & 0 & A-I \\ (A^2 - A + 2I)(I-A) & A(I-A) & (I-A)^2 \end{pmatrix} & k=1 \\ \begin{pmatrix} -A^{k-2}(I+A^2)(I-A)^2 & -A^{k-1}(I-A)^2 & (A^k - A^{k-1} + A^{k-2})(A-I) \\ -A^{k-1}(I-A)^2 & 0 & A^{k-1}(A-I) \\ (A^{k+1} - A^k + A^{k-1})(I-A) & A^k(I-A) & A^{k-1}(I-A)^2 \end{pmatrix} & k > 1 \end{cases}$$

Proof. The proof is straightforward from Proposition 1. \square

Lemma 5 directly leads to the following proposition which gives a necessary condition for the optimality of the OUT policy. We present it without proof.

Proposition 2. $\lambda^* = 0$ only if

$$\frac{\partial \Sigma_{ii}}{\partial \lambda} = Q \left(P_1 + 2 \sum_{k=1}^{L^+-1} \bar{\Psi}_L(k) P_{2k} + \sum_{k=1}^{L^+-2} \nu_k P_{3k} \right) \Delta Q^T = 0. \quad (21)$$

For all combinations of ARMA processes and discrete lead-time distributions, the set that satisfies (21) has zero measure (loosely speaking, for a randomly chosen demand model

and lead-time distribution, the probability that $\lambda^* = 0$ is negligible). That is, the demand process and lead-time distribution seldom meet the criteria of (21).

On the other hand, the signs of the derivatives at $\lambda = 0$ determine whether an optimal proportional policy leads to an increased or decreased order variance, as Proposition 3 shows.

Proposition 3. *If (19) and (20) have the same sign, then $\Sigma_{oo}|_{\lambda=\lambda^*} < \Sigma_{oo}|_{\lambda=0}$.*

We have omitted the proof because it is straightforward from Lemma 5. The logic is that if both derivatives are positive or negative at $\lambda = 0$, then decreasing Σ_{ii} will also decrease Σ_{oo} . The inverse proposition is also true. Note that Proposition 3 builds on the assumption that Σ_{oo} is monotonic for $\lambda \in (0, \lambda^*)$. This can be reasonably accepted since λ^* is usually quite close to 0.

Through the state space formulation, Propositions 2 and 3 provide computationally efficient criteria to determine the performance of OUT policy under ARMA demand and stochastic lead-time. For a demand process that is temporally independent, the non-optimality and non-existence of variance trade-off at $\lambda = 0$ is guaranteed when order crossover is present:

Corollary 1. *Under i.i.d. demand, MMSE forecast and order crossover, $\lambda^* \neq 0$ and $\Sigma_{oo}|_{\lambda=\lambda^*} < \Sigma_{oo}|_{\lambda=0}$.*

Corollary 1 contradicts Lemma 1 in the presence of order crossover. This is because all the inventory sub-processes characterized by the vector ξ , with an order crossover (when there is at least one 1 on the left side of a 0 in ξ), can all be minimized with a $\lambda > 0$, whereas those without crossover are minimized by $\lambda = 0$. Therefore the full inventory process, which is a sum of all the sub-processes, is minimized by a $\lambda > 0$. Intuitively, this is because under order crossover, the WIP is a weighted sum of multiple outstanding orders instead of simple sum. Hence the feedback controller of inventory position in the replenishment policy should be a fraction.

4. Numerical experiments

In this section we conduct numerical experiments to illustrate our analytical results. We show the mixed impact of the stochastic lead-time, order crossover and demand correlation. In all experiments the expected demand is 5 units per period. $\{\varepsilon_t\}$ is a white noise process, i.i.d. random variable with zero mean and unit variance. We consider 10 different lead-time distributions (see Figure 1) with both non-crossover (i, ii) and crossover (iii-x) scenarios. In the crossover scenarios we further investigate several distributions where the lead-time variability gradually increases, from low variability distributions (iii, vii), to higher variability (iv, viii), uniform (v, ix) and then two-point distributions (vi, x).

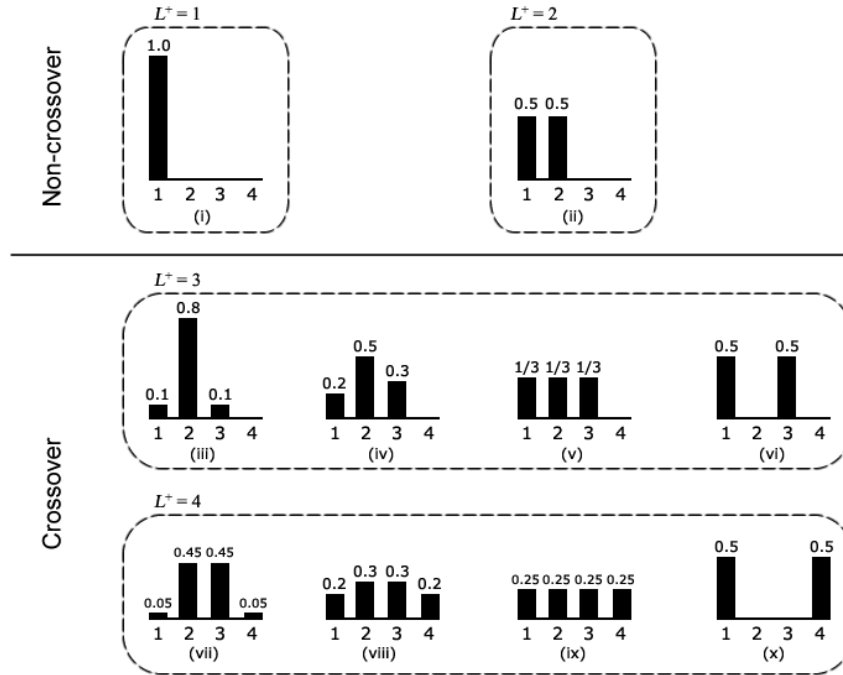


Figure 1: Lead-time distribution scenarios in the numerical experiment

4.1. The impact of stochastic lead-time

In Table 3 we compare the inventory and order variances under the OUT policy and the optimized POUT policy. Here demand follows an i.i.d. white noise process. First of all, we can observe that *the proportional policy can reduce order and inventory variance*

Table 3: System variance with different lead time distributions and i.i.d. demand

Lead-time distribution			OUT ($1 - \lambda = 1$)		POUT			% Σ_{ii}	% Σ_{oo}
No.	Description		Σ_{ii}	Σ_{oo}	$1 - \lambda^*$	Σ_{ii}	Σ_{oo}	Reduction	Reduction
i	$L^- = L^+ = 1$	$p_1 = 1$	1	1	1	1	1	0.00	0.00
ii	$L^- = 1, L^+ = 2$	$p_1 = 0.5, p_2 = 0.5$	7.75	1	1	7.75	1	0.00	0.00
iii	$L^- = 1, L^+ = 3$	$p_1 = 0.1, p_2 = 0.8, p_3 = 0.1$	6.50	1	0.99	6.50	0.98	0.00	2.00
iv		$p_1 = 0.2, p_2 = 0.5, p_3 = 0.3$	11.35	1	0.95	11.35	0.91	0.00	9.00
v		$p_l = 1/3, l = 1, 2, 3$	13.11	1	0.92	13.10	0.85	0.08	15.00
vi		$p_1 = 0.5, p_3 = 0.5$	14.50	1	0.87	14.47	0.76	0.21	24.00
vii		$p_1 = 0.05, p_2 = 0.45,$ $p_3 = 0.45, p_4 = 0.05$	11.12	1	0.96	11.12	0.92	0.00	8.00
viii	$L^- = 1, L^+ = 4$	$p_1 = 0.2, p_2 = 0.3,$ $p_3 = 0.3, p_4 = 0.2$	16.75	1	0.88	16.73	0.78	0.12	22.00
ix		$p_l = 1/4, l = 1, 2, 3, 4$	18.13	1	0.86	18.09	0.75	0.22	25.00
x		$p_1 = 0.5, p_4 = 0.5$	21.25	1	0.79	21.14	0.65	0.52	35.00

simultaneously – but only when there is order crossover. When crossover is not present (cases i and ii), $\lambda = 0$ minimizes inventory variance. For the order crossover scenarios (iii-x), an optimized POUT policy is able to mitigate the order amplification significantly while mildly reducing the inventory variance. The reduction in order variability is more considerable when lead-time volatility is high, e.g., when it follows a two-point distribution (cases vi and x). When the lead-time distribution is less variable, λ^* is close to zero, and the benefit of the POUT policy becomes marginal.

For illustration, we have plotted the inventory and order variances w.r.t. λ for case vi (see Figure 2) under i.i.d. demand. The existence of a minimal Σ_{ii} and the monotonicity of Σ_{oo} can be clearly seen. The inventory variance reaches a minimum when $\lambda = 0.13$ whereas the order variance decreases with λ . So setting λ to 0.13 reduces both inventory and order variances, compared with $\lambda = 0$. As λ increases, the order variance will continue to decrease; however this comes with increased inventory variance. (For instance, when $\lambda = 0.26$, the inventory variance is the same as when $\lambda = 0$; the order variance has reduced from 1 to 0.59.)

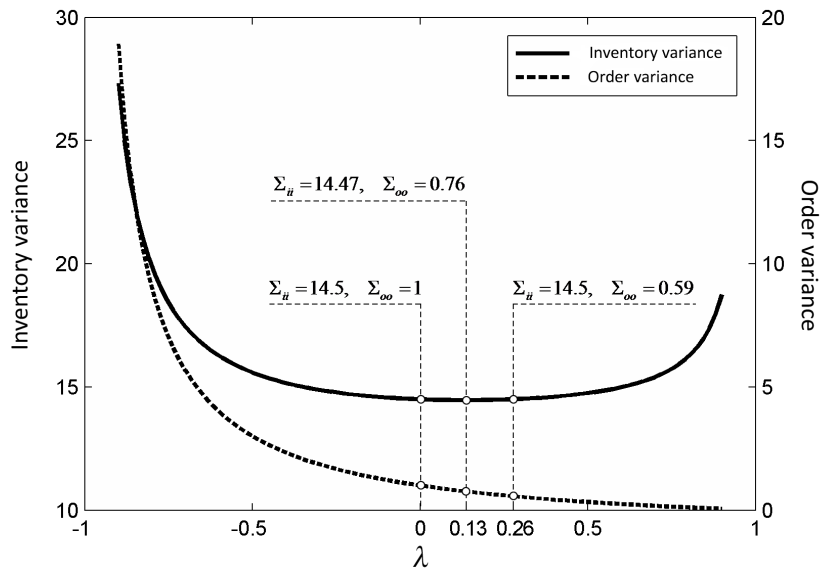


Figure 2: Inventory variance and bullwhip in Case vi

In Table 4 we make a similar comparison between the OUT and proportional policies, the only difference with Table 3 being that demand follows an AR(2) process with $\phi_1 = 0.6$ and $\phi_2 = -0.9$. Similar observations can be made.

The impact of the POUT policy on order and inventory variance is asymmetric. This can be explained as follows. Under stochastic lead-time and positive average demand, the inventory distribution is multimodal with modes separated by μ_d . While the POUT policy reduces the variance of each mode, it has no effect on the mode separation. Therefore we see that the inventory variance is dominated by μ_d rather than λ . The order distribution is unaffected by μ_d and is always a unimodal normal distribution; because of this, the influence of λ on order variance is more substantial.

4.2. Optimality and variance trade-off: Special cases

For demand processes following an AR(2) and ARMA(1,1) processes, we investigate the optimality of OUT policy and the possibility of reducing order and inventory variance simultaneously. The reason we have chosen these two processes is that the parametrical space is two dimensional which can be easily visualized. The AR(1) process is naturally

Table 4: System variance with different lead time distributions and AR(2) demand

Lead-time distribution			OUT ($1 - \lambda = 1$)		POUT			% Σ_{ii}	% Σ_{oo}
No.	Description		Σ_{ii}	Σ_{oo}	$1 - \lambda^*$	Σ_{ii}	Σ_{oo}	Reduction	Reduction
i	$L^- = L^+ = 1$	$p_1 = 1$	1	7.05	1	1	7.05	0.00	0.00
ii	$L^- = 1, L^+ = 2$	$p_1 = 0.5, p_2 = 0.5$	9.65	7.42	1	9.65	7.42	0.00	0.00
iii	$L^- = 1, L^+ = 3$	$p_1 = 0.1, p_2 = 0.8, p_3 = 0.1$	8.73	4.19	0.99	8.73	4.13	0.00	1.43
iv		$p_1 = 0.2, p_2 = 0.5, p_3 = 0.3$	14.43	2.64	0.94	14.42	2.43	0.07	7.95
v		$p_l = 1/3, l = 1, 2, 3$	16.50	2.16	0.91	16.48	1.87	0.12	13.43
vi		$p_1 = 0.5, p_3 = 0.5$	18.37	1.24	0.85	18.32	0.92	0.27	25.81
vii		$p_1 = 0.05, p_2 = 0.45,$ $p_3 = 0.45, p_4 = 0.05$	14.15	2.26	0.95	14.15	2.15	0.00	4.87
viii	$L^- = 1, L^+ = 4$	$p_1 = 0.2, p_2 = 0.3,$ $p_3 = 0.3, p_4 = 0.2$	20.51	1.05	0.86	20.48	0.83	0.15	20.95
ix		$p_l = 1/4, l = 1, 2, 3, 4$	21.98	0.83	0.85	21.94	0.60	0.18	27.71
x		$p_1 = 0.5, p_4 = 0.5$	24.45	1.13	0.79	24.42	0.94	0.12	16.81

included in the analysis as a special case. The lead-time distribution follows the cases in Figure 1, excluding (i) and (ii). Figures 3 and 4 show the auto-correlation and moving average parameters under which the OUT policy is optimal (bold curve) and where the variance trade-off exists (dark grey area). The white area denotes the desirable situation where minimizing the inventory variance simultaneously results in a reduction in the order variance. The dashed line for $\phi_2 = 0$ and $\theta = 0$ represents the AR(1) case. The light grey area in Figure 3 is the unstable region for AR(2) process.

First we note, the condition of OUT optimality (21) is represented by one-dimensional curves in the figures, therefore it is practically impossible for the OUT policy to be optimal. Secondly, as anticipated in §3.4, the area where the variance trade-off exists for the OUT policy only takes a small proportion of the parametrical plane. In some cases (iii and vii under ARMA(1,1) demand) such an area does not exist at all. Thirdly, in most cases, OUT optimality and variance trade-off usually occurs when dominant auto-regressive parameter is negative ($\phi_1 < 0$).

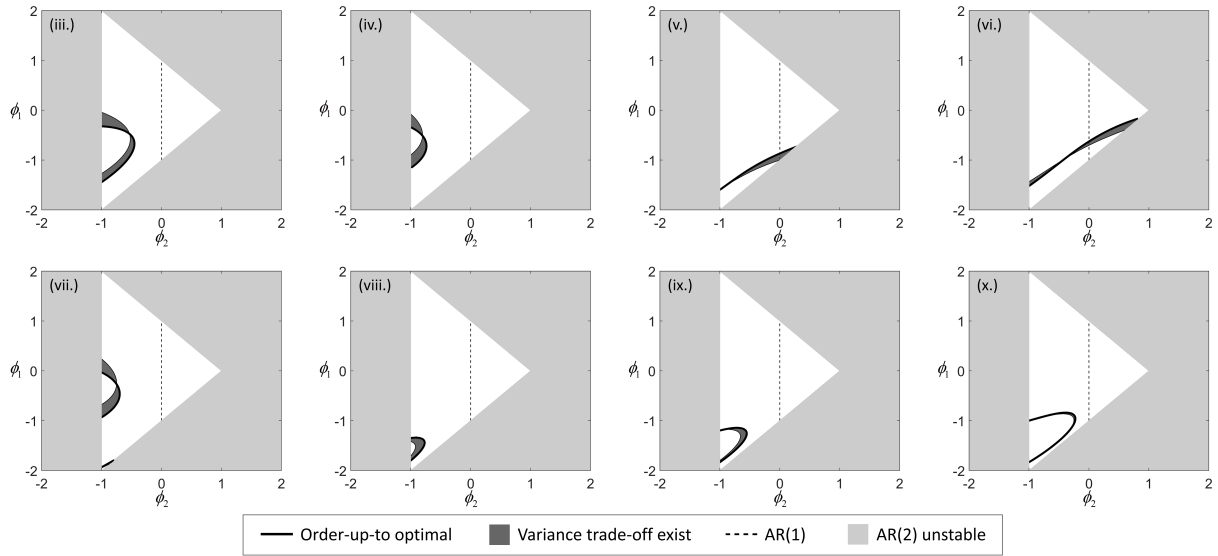


Figure 3: Optimality and trade-off of order-up-to policy under AR(2) process and stochastic lead-time

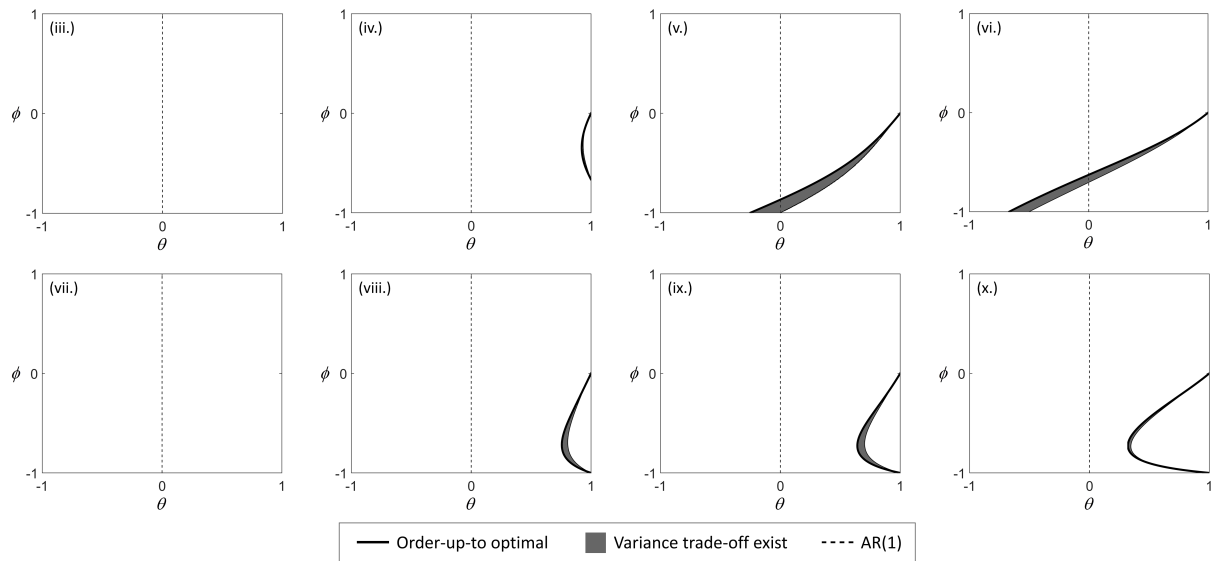


Figure 4: Optimality and trade-off of order-up-to policy under ARMA(1,1) process and stochastic lead-time

4.3. Cost implications using real demand data

In this section we use real demand data and lead-time distributions to test the performance of the proportional policy, in terms of both variance and cost. The dataset contains time series of demand from a general household retailer in Germany. Each time series con-

tain 103 data points. The ARMA(1,1) process is found to be an appropriate model for 91 items in the dataset, and Figure 5 portrays the ARMA(1,1) coefficients, ϕ and θ for these items. The empirical data shown in Figure 2 of Disney et al. (2016) is used to define the lead-time distribution. In Figure 5 we present the optimality and trade-off details under this demand distribution together with the auto-correlation and moving average parameters of the demand processes for each of the 91 ARMA(1,1) items. It can be clearly seen that the POUT policy should be able to mitigate inventory and order fluctuation for all of the items in the dataset.

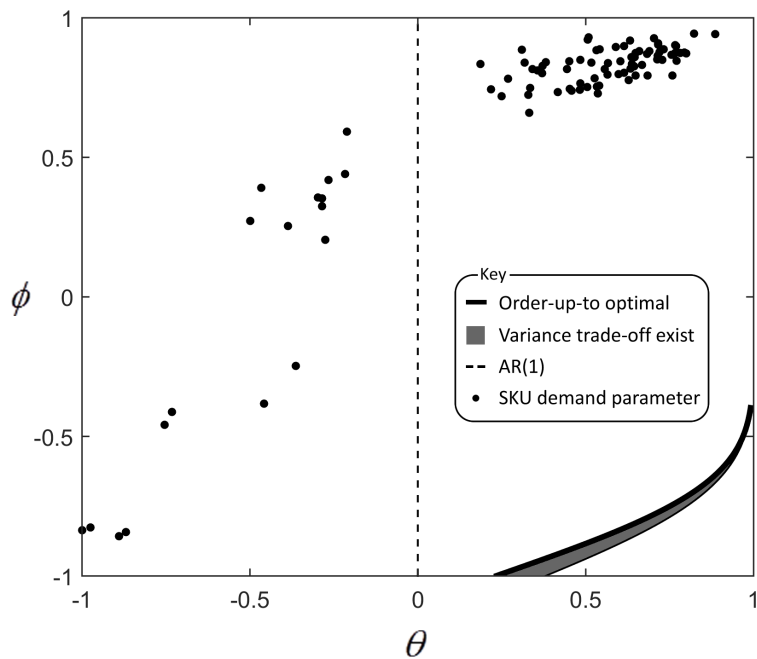


Figure 5: Optimality and trade-off under real lead-time distribution

We use simulation to verify this result. First we derive the optimal feedback controller for each item based on model fit. The demand and lead-time data is then imported into the inventory system which uses either the OUT policy or the optimized POUT policy. The variance ratio between inventory/order and demand is recorded. To capture the effect of the randomly generated lead-time, the simulation is repeated 100 times with different lead-time realizations in each repetition and the average is taken. Only the last 50 data points are

Table 5: Average variance ratio with real demand data and lead-time distribution

Replenishment policy	Σ_{ii}/Σ_{dd}	Σ_{oo}/Σ_{dd}	% Reduction Σ_{ii}/Σ_{dd}	% Reduction Σ_{oo}/Σ_{dd}
OUT	29.854	16.949	-	-
POUT	28.947	14.903	3.04*	12.07*

* The difference is significant at the 0.05 level.

used in the performance evaluation to exclude any initialization effects in our simulation. We present the result both in the form of average over the items (Table 5) and a boxplot of all items (Figure 6).

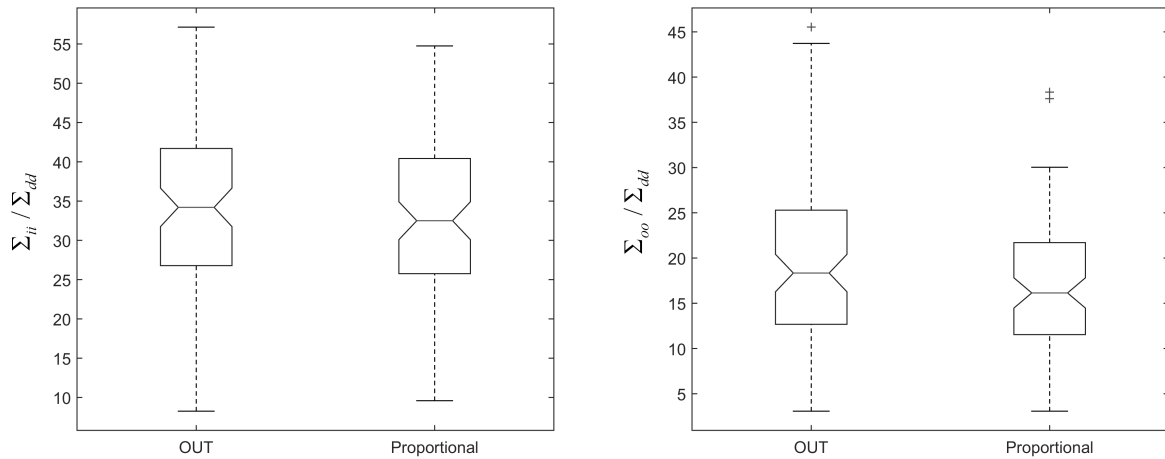


Figure 6: Boxplot of variance ratio for all SKUs (The range of whiskers is $\pm 2.7\sigma$).

The simulated performance is consistent with our analytical results. By adopting the proportional policy, the inventory variance can be reduced mildly (3.04%). However the order variance can be reduced significantly (12.07%). This result verifies the effectiveness of proportional inventory control in the presence of a stochastic lead-time. From the boxplot it can be seen that the proportional policy also leads to a decrease in the variation among items.

The cost performances of the order-up-to and proportional policies are also compared.

Table 6: Inventory cost with real demand data and lead-time distribution under various cost parameters

Replenishment policy	$h = 5, b = 5$		$h = 3, b = 7$		$h = 1, b = 9$	
	C	% Reduction	C	% Reduction	C	% Reduction
OUT	587.77	-	514.42	-	265.62	-
POUT	579.20	1.46	507.13	1.42	263.43	0.82

In Table 6 the inventory cost as defined by (2). The safety stock is set via $\int_{-\infty}^0 \psi_i(x) dx = h/(h+b)$ in the usual newsvendor fashion. Three cost settings are included: $\{h = 5, b = 5\}$, $\{h = 3, b = 7\}$ and $\{h = 1, b = 9\}$. They represent different availability requirements of 50%, 70% and 90% respectively. The numerical results confirm that the proportional policy also reduces average inventory cost, especially when the service level is low.

5. Conclusion and discussion

Based on our experience of observing a stochastic lead-time in practice and building upon the results in Disney et al. (2016), we studied its effect on order and inventory variances with auto-correlated demand. Given a known (or perceived) lead-time distribution and demand correlation information, we presented a state-space approach to calculate the inventory variance. We gave conditions under which the POUT policy outperforms the OUT policy in mitigating both order and inventory amplifications.

In Table 7 we summarize the properties of the inventory system under different demand and lead-time conditions. First, the stochasticity of the lead-time determines the shape of the inventory distribution, and hence the relationship between demand level and inventory variance. If the lead-time is stochastic, then the inventory variance is dependent upon the average demand level. Second, order crossover determines whether the OUT policy is the proper choice for minimizing inventory variance and inventory related costs. Proportional control is able to reduce inventory variance under order crossover. Third, the demand correlation corresponds to whether there is a variance trade-off under the OUT policy. An investigation of the special cases of AR(2) and ARMA(1,1) demand suggests that order and

Table 7: Summarization of the properties of different lead-time conditions

	Constant lead-time, iid/ARMA demand	Stochastic lead-time without crossover, iid demand	Order crossover, i.i.d. demand	Order crossover, ARMA demand
Related research	Disney et al. (2004), Gaalman (2006)	Disney et al. (2016), this paper	Disney et al. (2016), this paper	This paper
Shape of inventory pdf	Bell-shaped	Multi-modal		
μ_d and Σ_{ii}	Σ_{ii} is independent of μ_d	Σ_{ii} is dependent on μ_d		
λ^*	$\lambda^* = 0$		$\lambda^* \neq 0$	$\lambda^* \neq 0$ in most cases
Variance trade-off	Reducing Σ_{oo} increases Σ_{ii}		Minimizing Σ_{ii} decreases Σ_{oo}	Depending on demand correlation and lead-time distribution

inventory variance can be almost always be simultaneously reduced.

Managerially this result has interesting consequences. One of the concerns practitioners have before adopting a POUT policy is its potential aggravating effect on inventory cost, see Lemma 1. However, our finding implies that, under most demand correlation and order crossover scenarios (which are quite common in practice), this concern is not valid since inventory variance and inventory cost can be reduced by proportional control, albeit only marginally. However, this subsequently allows one to mitigate order fluctuation considerably. Even if the proportional controller is not set optimally due to statistical and computational imperfections, the benefit of reducing order variability will arguably outweigh the (small) potential loss of the increased inventory cost. The intuition is that *the order variability is always more sensitive to the proportional controller, and inventory cost is less sensitive, especially when the mean demand is large and the lead-time is highly variable*. Therefore we conclude that the proportional policy is preferable to the OUT policy in the presence of volatile delays and auto-correlated demand, conditions which co-exist in contemporary global supply chains.

This research can be further developed in the following directions. (1) The model assumes full knowledge of the demand and lead-time distributions. Although this can be

partly achieved by statistical analysis of historical data, a dynamic inventory policy that tracks this data over time might be more desirable. (2) An exogenous lead-time distribution is assumed throughout this paper for simplicity. However, in practice we may be able to adjust the transportation mode (and hence the lead-time) based on system states. For example in times of low stock or high demand, air freight could be used to expedite replenishment. Thus it will be intriguing to incorporate dependence of transportation delay on system states. (3) The ARMA(p,q) model can be further extended to ARIMA(p,d,q) model or seasonal ARMA(p,q) model (Nagaraja et al., 2015) to incorporate more sophisticated demand patterns; (4) From the theoretical perspective, this problem calls for more knowledge on mathematical properties of the objectives, e.g. monotonicity and convexity of order and inventory variances w.r.t. λ for ARMA demand and stochastic lead-time.

Acknowledgments

We would like to thank Professor Aris Syntetos for providing the demand dataset in §4.3.

References

- Agin, N., 1966. A min-max inventory model. *Management Science* 12, 517–529.
- Bagchi, U., Hayya, J., Chu, C.-H., 1986. The effect of lead time variability: the case of independent demand. *Journal of Operations Management* 6, 159–177.
- Bischof, D., Robb, D., Silver, E., J.D., B., 2014. Analysis and management of periodic review, order-up-to level inventory systems with order crossover. *Production and Operations Management* 23 (5), 762–772.
- Bradley, J., Robinson, L., 2005. Improved base-stock approximations for independent stochastic lead times with order crossover. *Manufacturing and Service Operations Management* 7, 319–329.
- Chatfield, D., Kim, J., Harrison, T., Hayya, J., 2004. The bullwhip effect - impact of stochastic lead time, information quality, and information sharing: a simulation study. *Production and Operations Management* 13, 340–353.
- Chen, F., Dresner, Z., Ryan, J., Simchi-Levi, D., 2000. Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science* 46, 436–443.
- Chen, Y., Disney, S., 2007. The myopic order-up-to policy with a proportional feedback controller. *International Journal of Production Research* 45, 351–368.

- Dejonckheere, J., Disney, S., Lambrecht, M., Towill, D., 2003. Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operational Research* 147 (3), 567–590.
- Deziel, D., Eilon, S., 1967. A linear production-inventory control rule. *The Production Engineer* 43, 93–104.
- Disney, S., 2008. Supply chain aperiodicity, bullwhip and stability analysis with jurys inners. *IMA Journal of Management Mathematics* 19 (2), 101–116.
- Disney, S., Maltz, A., Wang, X., Warburton, R., 2016. Inventory management for stochastic lead times with order crossovers. *European Journal of Operational Research* 248, 473–486.
- Disney, S., Towill, D., Van de Velde, W., 2004. Variance amplification and the golden ratio in production and inventory control. *International Journal of Production Economics* 90, 295–309.
- Finch, P., 1961. The order cycle system of stock control. *Journal of the Australian Mathematical Society* 2, 209–216.
- Gaalman, G., 2006. Bullwhip reduction for ARMA demand: The proportional order-up-to policy versus the full-state-feedback policy. *Automatica* 42, 1283–1290.
- Hadley, G., Whitin, T., 1963. *Analysis of Inventory Systems*. Prentice-Hall, Englewood Cliffs, NJ.
- Hayya, J., Bagchi, U., Kim, J., Sun, D., 2008a. On static stochastic order crossover. *International Journal of Production Research* 114, 404–413.
- Hayya, J., Harrison, T., Chatfield, D., 2008b. A solution for the intractable inventory model when both demand and lead time are stochastic. *International Journal of Production Economics* 122, 595–605.
- Hayya, J., Harrison, T., He, H., 2011. The impact of stochastic lead time reduction on inventory cost under order crossover. *European Journal of Operational Research* 211, 274–281.
- Hosoda, T., Disney, S., 2006. On variance amplification in a three-echelon supply chain with minimum mean square error forecasting. *Omega* 34 (2), 344–358.
- John, S., Naim, M., Towill, D., 1994. Dynamic analysis of a wip compensated decision support system. *International Journal of Manufacturing System Design* 1 (4), 283–297.
- Kim, J., Chatfield, D., Harrison, T., Hayya, J., 2006. Quantifying the bullwhip effect in a supply chain with stochastic lead time. *European Journal of Operational Research* 173, 617–636.
- Lee, H., Padmanabhan, V., Whang, S., 1997. Information distortion in a supply chain: The bullwhip effect. *Management Science* 43, 546–558.
- Magee, J., 1956. Guides to inventory control part II. *Harvard Business Review* 34, 103–116.
- Nagaraja, C., Thavaneswaran, A., Appadoo, S., 2015. Measuring the bullwhip effect for supply chains with seasonal demand components. *European Journal of Operational Research* 242 (2), 445–454.
- Riezebos, J., 2006. Inventory order crossovers. *International Journal of Production Economics* 104, 666–675.
- Robinson, L., Bradley, J., Thomas, L., 2001. Consequences of order crossover under order-up-to inventory policies. *Manufacturing and Service Operations Management* 3, 175–188.

Song, J.-S., 1994. The effect of lead time uncertainty in a simple stochastic inventory model. *Management Science* 40, 603–613.

Zalkind, D., 1978. Order-level inventory systems with independent stochastic lead times. *Management Science* 24, 1384–1392.

Zipkin, P., 1986. Stochastic lead times in continuous-time inventory models. *Naval Research Logistics Quarterly* 33, 763–774.

Appendix A. Mathematical Proofs

Proof of Lemma 2. We prove the normality of order distribution by showing that under a stochastic lead-time, the order process is the output of a linear system with a normally distributed input. The inventory system can be defined by two difference equations: $o_t = f_t + (1 - \lambda)(ss - IP_t)$ and $IP_t = IP_{t-1} + o_{t-1} - d_t$. Both d_t and f_t are linear functions of y_t , which is normally distributed. So o_t and IP_t are also normally distributed.

$E(o)$ can be derived directly from $IP_t = IP_{t-1} + o_{t-1} - d_t$ by taking expectations on both sides and rearranging. From (11), $E(f) = (\mu_L - \lambda\mu_L + \lambda)\mu_d$. Together with (10) we have the mean of inventory position. \square

Proof of Lemma 3. Substituting (1) into (10) and arranging, we have

$$o_t = f_t - f_{t-1} + \lambda o_{t-1} + (1 - \lambda)d_t$$

which can be further rewritten as

$$o_t = [FA + (1 - \lambda)CA - F]y_{t-1} + \lambda o_{t-1} + [F + (1 - \lambda)C]B\varepsilon_t + (1 - \lambda)\mu_d. \quad (\text{A.1})$$

To calculate the correlation function of x and y we simply multiply x by the transpose of y and take the expectation. Note that the constants can be omitted when they do not contribute to second order moments. Therefore we multiply (A.1) by $o_{t-\tau}$ on both sides and take expectation to obtain (15). Using the same method, multiply $o_{t-\tau}$ on both sides of $y_t = Ay_{t-1} + B\varepsilon_t$ and notice ε_t is independent in time, we have (13). The initial values of the iterations (15) and (13) can be calculated in the following way. Firstly, y and o form a

stochastic affine system:

$$\begin{pmatrix} y_t \\ o_t \end{pmatrix} = \begin{pmatrix} A & 0 \\ FA + (1 - \lambda)CA - F & \lambda \end{pmatrix} \begin{pmatrix} y_{t-1} \\ o_{t-1} \end{pmatrix} + \begin{pmatrix} B \\ FB + (1 - \lambda)CB \end{pmatrix} \varepsilon_t.$$

Multiplying both sides with transposition shows that the covariance of the state variables satisfies

$$\begin{pmatrix} \Sigma_{yy} & \Sigma_{yo} \\ \Sigma_{oy} & \sigma_{oo} \end{pmatrix} = \begin{pmatrix} A & 0 \\ FA + (1 - \lambda)CA - F & \lambda \end{pmatrix} \begin{pmatrix} \Sigma_{yy} & \Sigma_{yo} \\ \Sigma_{oy} & \sigma_{oo} \end{pmatrix} \begin{pmatrix} A & 0 \\ FA + (1 - \lambda)CA - F & \lambda \end{pmatrix}^T + \begin{pmatrix} B \\ FB + (1 - \lambda)CB \end{pmatrix} \begin{pmatrix} B \\ FB + (1 - \lambda)CB \end{pmatrix}^T.$$

The top right element gives (12), and the bottom right element gives (14). \square

Proof of Lemma 4. For the mean of each inventory sub-process, note that

$$E(w; \xi) = \sum_{k=1}^{L^+ - 1} \xi(k) \mu_d = \xi \mathbf{1} \mu_d.$$

By Lemma 2 we see that the mean of inventory position is independent of the lead-time. Therefore $E(i; \xi) = E(IP) - E(w; \xi)$. For the variance of a specific inventory sub-process, rearrange (10) to see that

$$i_t = \frac{f_t - o_t}{1 - \lambda} - w_t + ss.$$

The variance of an inventory sub-process is then

$$\Sigma_{ii}(0; \xi) = \frac{\Sigma_{ff} + \Sigma_{oo} - 2F\Sigma_{yo}}{(1 - \lambda)^2} + 2 \frac{\Sigma_{ow}(0; \xi) - \Sigma_{fw}(0; \xi)}{1 - \lambda} + \Sigma_{ww}(0; \xi)$$

where $\Sigma_{fo} = F\Sigma_{yo}$. The covariances $\Sigma_{ow}(0; \xi)$, $\Sigma_{fw}(0; \xi)$ and $\Sigma_{ww}(0; \xi)$ can be expressed in the matrix form: $\Sigma_{ow}(0; \xi) = \xi\alpha$, $\Sigma_{fw}(0; \xi) = \xi\beta$ and $\Sigma_{ww}(0; \xi) = \xi\Gamma\xi^T$ due to their linearity. \square

Proof of Proposition 1. The distribution of the inventory is a weighted sum of normal distributions. Therefore the variance of the inventory process can be calculated as follows:

$$\Sigma_{ii} = \sum_{j=1}^{2^{L^+ - L^-}} p(\xi^j) \Sigma_{ii}(0; \xi^j) + \frac{1}{2} \sum_{j,k=1}^{2^{L^+ - L^-}} p(\xi^j) p(\xi^k) [E(i; \xi^j) - E(i; \xi^k)]^2.$$

The second term leads to the last term in (18). Note that Σ_{NN} is the variance of the number of outstanding orders.

By the definition of $p(\xi)$ and WIP we see that $\sum_k p(\xi^k) [\Sigma_{ow}(0; \xi^k) - \Sigma_{fw}(0; \xi^k)]$ can be interpreted as the summation of $\Sigma_{oo}(k) - \Sigma_{fo}(k)$ multiplied by the expected number of outstanding orders in the pipeline, and $\sum_k p(\xi^k) \Sigma_{ww}(0; \xi^k)$ as the summation of $\Sigma_{oo}(k)$ multiplied by the expected number of order pairs in the pipeline. Thus

$$\sum_{k=1}^{2^{L^+ - L^-}} p(\xi^k) \xi^k (\alpha - \beta) = \sum_{k=1}^{L^+ - 1} \bar{\Psi}_L(k) [\Sigma_{oo}(k) - F \Sigma_{yo}(k)]$$

and

$$\sum_{k=1}^{2^{L^+ - L^-}} p(\xi^k) \xi^k \Gamma(\xi^k)^T = \sum_{k=1}^{L^+ - 1} \bar{\Psi}_L(k) \Sigma_{oo} + \sum_{k=1}^{L^+ - 2} \nu_k \Sigma_{oo}(k).$$

$\bar{\Psi}_L(k)$ is the probability that an order placed k periods ago is still outstanding; ν_k is the probability that two orders with k periods gap in placement are both outstanding, which is computed as stated in Proposition 1. Substitution and rearranging will lead us to (18). \square

Proof of Corollary 1. We complete this proof by showing that under i.i.d. demand, $\partial \Sigma_{ii} / \partial \lambda < 0$ when $\lambda = 0$, indicating that there exists a $\lambda > 0$ that outperforms $\lambda = 0$. As demand is i.i.d. and a MMSE forecast is adopted, we have $A = 0$, $F = 0$, $C = 1$, $\Delta = I$. Also $P_1 I = 0$, $QP_{2k}Q^T = \begin{cases} 0 & k = 1 \\ -1 & k > 1 \end{cases}$, and $QP_{3k}Q^T = \begin{cases} 1 & k = 1 \\ 0 & k > 1 \end{cases}$. Then (21) becomes

$$\begin{aligned} \nu_1 - 2 \sum_{k=2}^{L^+ - 1} \bar{\Psi}_L(k) &= 2 \sum_{k=2}^{L^+ - 1} \bar{\Psi}_L(k-1) \bar{\Psi}_L(k) - 2 \sum_{k=2}^{L^+ - 1} \bar{\Psi}_L(k) \\ &= 2 \sum_{k=2}^{L^+ - 1} \bar{\Psi}_L(k) [\bar{\Psi}_L(k-1) - 1]. \end{aligned}$$

With i.i.d. lead-times, the non-crossover distribution must satisfy $L^+ - L^- \leq 1$, which can be derived directly from Corollaries 1 and 2 in Riezebos (2006) and the exogenous assumption in this paper. Since $\bar{\Psi}_L(k)$ is non-increasing, the above equation equals to zero only in two cases: either $\bar{\Psi}_L(L^+ - 1) = 1$, which means that the lead-time is constant; or $\bar{\Psi}_L(L^+ - 1) < 1$ and $\bar{\Psi}_L(L^+ - 2) = 1$, which means the lead-time is stochastic but there is

no order crossover. Otherwise $\partial \Sigma_{ii} / \partial \lambda|_{\lambda=0} < 0$, from which we also have $\lambda^* > 0$. Since Σ_{oo} is monotonic and decreasing from Lemma 1, we finally have $\Sigma_{oo}|_{\lambda=\lambda^*} < \Sigma_{oo}|_{\lambda=0}$. \square