On the order-up-to policy with intermittent integer demand and coherent forecasts

Bahman Rostami-Tabar[a], Stephen M. Disney[b,c]

[a] *Logistics Systems Dynamics Group, Cardiff Business School*
Cardiff University, Cardiff, CF10 3EU, United Kingdom
rostami-tabarb@cardiff.ac.uk
[b] *Centre for Simulation, Analytics, and Modelling, University of Exeter Business School,*
Streatham Court, Rennes Drive, Exeter, EX4 4PU, United Kingdom s.m.disney@exeter.ac.uk
[c] *Presenting author*

Abstract

We measure the impact of a first-order integer auto-regressive, INAR(1), demand process on order-up-to (OUT) replenishment policy dynamics. We obtain a unique understanding of the bullwhip behaviour for slow moving integer demand. We forecast this integer demand in two ways; with a *conditional mean* and a *conditional median*. We investigate the impact of the two forecasting methods on the bullwhip effect and inventory variance generated by the OUT replenishment policy. While the conditional mean forecasts result in the tightest inventory control, they result in real valued orders and inventory levels which is incoherent with the integer demand. However, the conditional median forecasts are integer-valued and produce coherent integer order and inventory levels. The conditional median forecasts minimise the expected absolute forecasting error, but it is not possible to obtain closed forms for the variances. Numerical experiments reveal existing results remain valid with high volume correlated demand. However, for low volume demand, the impact of the integer demand on the bullwhip effect is often significant. Conditional median bullwhip can be both lower and higher than the conditional mean bullwhip; indeed it can even be higher than a known conditional mean upper bound (that is valid for all lead times under real-valued, first-order auto-regressive, AR(1), demand), depending on the auto-regressive parameter. Numerical experiments confirm the conditional mean inventory variance is a lower bound for the conditional median inventory variance.

Keywords: Integer auto-regressive demand processes, Intermittent demand, Bullwhip effect, Conditional mean forecasts, Conditional median forecasts, Poisson distribution.

1. Introduction

The bullwhip effect refers to the tendency for supply chain replenishment decisions to amplify the variability of the demand when releasing production orders onto the factory shop floor or placing replenishment orders onto suppliers (Lee et al., 2000). A rich literature has been developed in the last 25 years on the bullwhip effect since the seminal work of Lee et al. (1997). Many of the existing studies on the bullwhip effect (Chen et al., 2000a; Dejonckheere et al., 2003) have assumed that real-valued demand exists. That is, demand can take on any number, even fractional values. For some products sold by volume or weight (for example, powders, granules, or liquids) this may

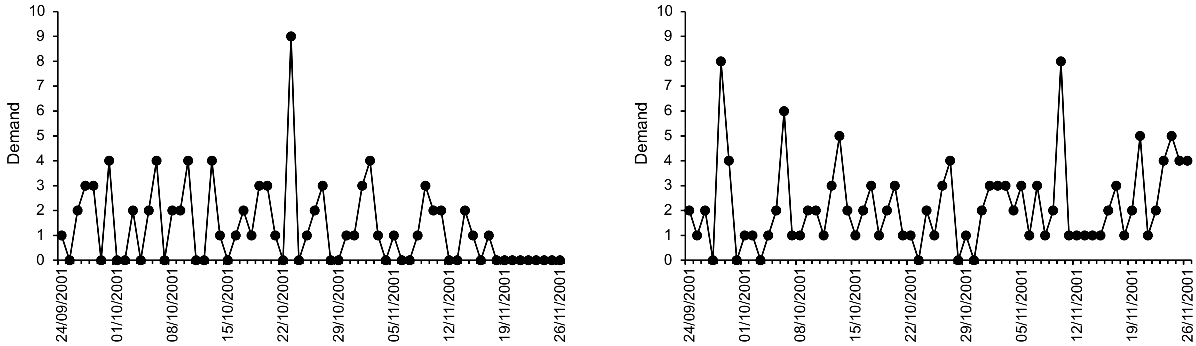

Figure 1: Low volume, occasionally zero, integer demand for two products from a single store of a UK grocery retailer

be appropriate. However, for other products, only integer (or batch) demand makes sense. For example, you can't buy half a bicycle or half a laptop. In these situations, demand and replenishment orders must be integers. For some situations, with high volume demand, replenishment calculations can ignore integer effects as rounding becomes negligible. However, low volume demand settings may be more susceptible to integer effects. Consider, for example, the daily demand for a single product in a single grocery store in Figure 1. These products have low volume, integer, occasionally zero, demands.

Continuous-valued time series can be modeled using auto-regressive integrated moving average (ARIMA) type processes (Box et al., 2015). First order ARIMA demand processes are popular assumptions in bullwhip studies due to their mathematical simplicity and their relevance to practical settings. These demand processes are often forecasted with minimum mean squared error (MMSE) forecasting methods (Graves, 1999; Chen et al., 2000a,b; Hosoda and Disney, 2006; Duc et al., 2008). However, standard ARIMA modeling techniques are not suitable for modelling non-negative integer-valued series (Silva and Oliveira, 2004). Therefore, another family of stationary models, integer auto-regressive moving average processes (INARMA), have been proposed, Al-Osh and Alzaid (1988).

Wang and Disney (2016) provide a comprehensive review of bullwhip research categorizing it into empirical, experimental, and analytical approaches. They note integer demand processes have not been extensively covered. The bullwhip effect can be measured in different ways, however it is often convenient to quantify it via the ratio of the variance of orders to the variance of demand. For normally distributed errors in the demand process, the capacity costs are known to be a linear function of the standard deviation of the orders. The standard deviation of the orders is closely related to the bullwhip measure. The other side of the bullwhip problem is the variance of the *inventory* levels. In a linear system (where a positive net stock represents inventory holding and a negative net stock represents unmet demand, or a backlog) and it is often convenient to use the term *net stock amplification* (*NSAmp*) as a moniker for the variance of the net stock divided by the variance of the demand. *NSAmp* is an important measure as it is closely related to the the expected, per period, inventory holding and backlog costs when normally distributed error terms are present in the demand process when a linear system system exists. Later in this paper we will explore the economic consequences of Poisson distributed errors and non-normal orders and net stock levels.

Our contribution is to investigate the consequences of integer demand on the performance of the OUT replenishment policy when two different forecasting approaches are used. The first forecasting method is based on the–frequently assumed–conditional mean of future demand (Wang and Disney, 2016). While, this forecasting method creates a MMSE forecast of future demands, it results in non-integer, forecasts, orders, and inventory levels. This is in tension with the integer demand assumption. The second forecasting method is based on the conditional median forecast which produces–conceptually consistent–integer forecasts, orders, and inventory levels. We find the conditional mean forecasts for INAR(1) demand results in exactly the same variance ratios as the conditional mean forecasts under real-valued AR(1) demand. The conditional median forecasts produce different order and inventory dynamics. Numerical experiments show existing results for real-valued demand can be used with confidence when we have high volume demand integer demand. However, under low volume, possibly intermittent, integer demand there can be a significant difference between the real- and integer-valued *Bullwhip* and *NSAmp* predictions.

The structure of the paper is as follows. §2 provides background information on the INAR(1) process and the OUT replenishment policy. §3 is devoted to creating conditional mean forecasts; §4 to creating conditional median forecasts. §5 characterises the *Bullwhip* and *NSAmp* performance of the OUT policy with the two different forecasts. Integer, independent and identically distributed (i.i.d.) demands are considered in more detail in §6. §7 concludes and highlights future research.

2. Model development

In this section, we first present the INAR(1) demand process and the OUT policy.

2.1. The INAR(1) demand process

We assume the demand follows a first-order integer auto-regressive process, INAR(1), where demand d in period t is given by:

$$d_t = \phi \circ d_{t-1} + z_t, \tag{1}$$

Here, d_t is the demand in period t, $0 \leq \phi \leq 1$ is the auto-regressive parameter, and z_t is a sequence of i.i.d. non-negative integer-valued Poisson distributed random variables, with mean λ and finite variance λ (Silva et al., 2009). The atomic expression $\phi \circ d_{t-1}$ is the binomial thinning operation,

$$\phi \circ d_{t-1} = \sum_{i=1}^{d_{t-1}} X_i. \tag{2}$$

Here, X_i is a sequence of i.i.d. Bernoulli indicators with parameter ϕ (i.e. with $\mathbb{P}(X_i = 1) = \phi$ for $i = \{1, 2, ..., d_{t-1}\}$). A natural interpretation of (1) is that d_t is the total number of guests in a hotel at time t, z_t is the number of new guests that arrived today, and $\phi \circ d_{t-1}$ is the number of guests that remained in the hotel from the day before (Ristić and Nastić, 2012). Silva and Oliveira (2004) provide a number of useful relations and properties of the INAR(1) model. Notably, the relations

$$\mathbb{E}[\phi \circ d_t] = \phi \mathbb{E}[d_t], \tag{3}$$

$$\mathbb{E}[\phi \circ d_t]^2 = \phi^2 \mathbb{E}[d_t^2] + \phi(1-\phi)\mathbb{E}[d_t], \text{ and} \tag{4}$$

$$\mathbb{V}[\phi \circ d_t] = \mathbb{E}[\phi \circ d_t]^2 - (\mathbb{E}[\phi \circ d_t])^2 = \phi \mathbb{V}[d_t], \tag{5}$$

are useful. Here, $\mathbb{E}[\cdot]$ is the expectation operator and $\mathbb{V}[\cdot]$ is the variance operator.

Lemma 1. *The INAR(1) demand process has mean, μ_d, and an auto-covariance with lag j, γ_j, of*

$$\mu_d = \frac{\lambda}{1-\phi} \quad \text{and} \quad \gamma_j = \begin{cases} \frac{\lambda}{1-\phi}, & j = 0, \\ \phi^j \gamma_0, & j \geq 1. \end{cases} \tag{6}$$

Proof. *The proof is provided in Appendix A.* □

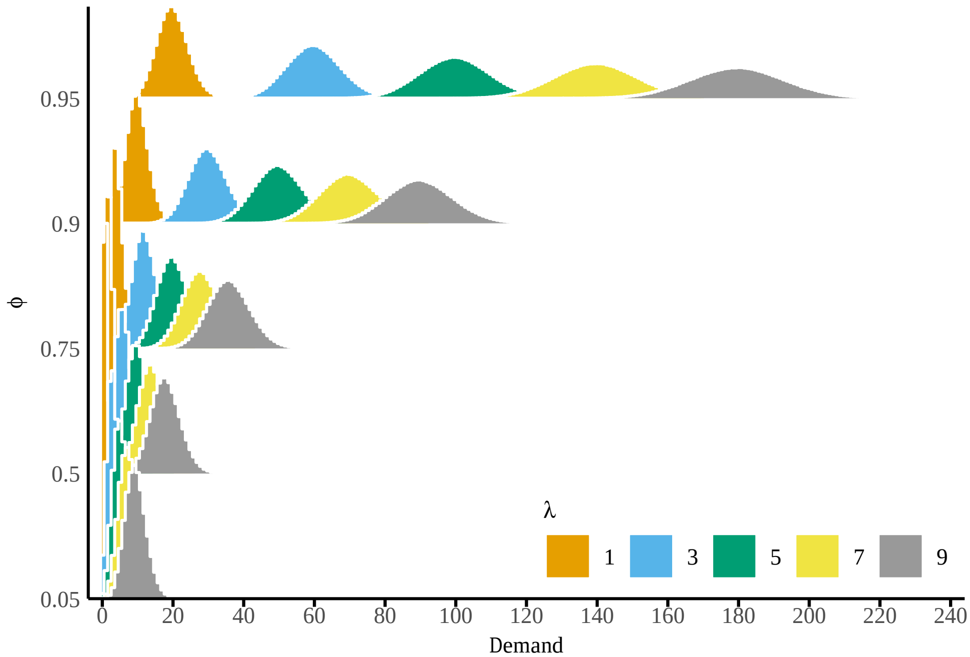

Figure 2: Probability mass function of the demand for different ϕ when $\lambda = 1$.

The auto-covariance function at lag $j = 0$ gives the demand variance. Silva et al. (2009) show the INAR(1) demand has a Poisson distribution with a shape parameter of $\lambda_d = \lambda/(1-\phi)$:

$$\mathbb{P}[d = k] = \frac{\lambda_d^k e^{-\lambda_d}}{k!}. \tag{7}$$

Figure 2 illustrates the probability mass function (pmf) of the demand for different ϕ and λ. The probability of a zero demand is increased with a small ϕ and small λ, demonstrating the power of the INAR(1) demand model for modelling low volume, intermittent, integer demand processes. Figure 3 illustrates example INAR(1) time series with different ϕ and λ in the right column of panels. When ϕ and λ are small the time series is intermittent; increasing ϕ (or λ, see (6)) increases the average demand. The left column of panels of Figure 3 provides an exact pmf (from (7)) of the corresponding time series. The probability of zero demand, $e^{-\lambda/(1-\phi)} \to 1$ when $\{\phi, \lambda\} \to 0$.

2.2. The order-up-to replenishment policy

At the end of period t, the retailer uses the OUT policy to order q_t items from the manufacturer;

$$q_t = s_t - s_{t-1} + d_t. \tag{8}$$

The order-up-to level, s_t, in time period t is determined by

$$s_t = \hat{d}_{t,L} + SS, \tag{9}$$

where $\hat{d}_{t,L} = \sum_{i=1}^{L} \hat{d}_{t+i}$ is the forecast of the demand over the lead-time L, SS is a safety stock used to achieve a given target level of inventory availability. The usually assumed, per period inventory holding and backlog costs

$$C_{i,t} = h[i_t]^+ + b[-i_t]^+, \tag{10}$$

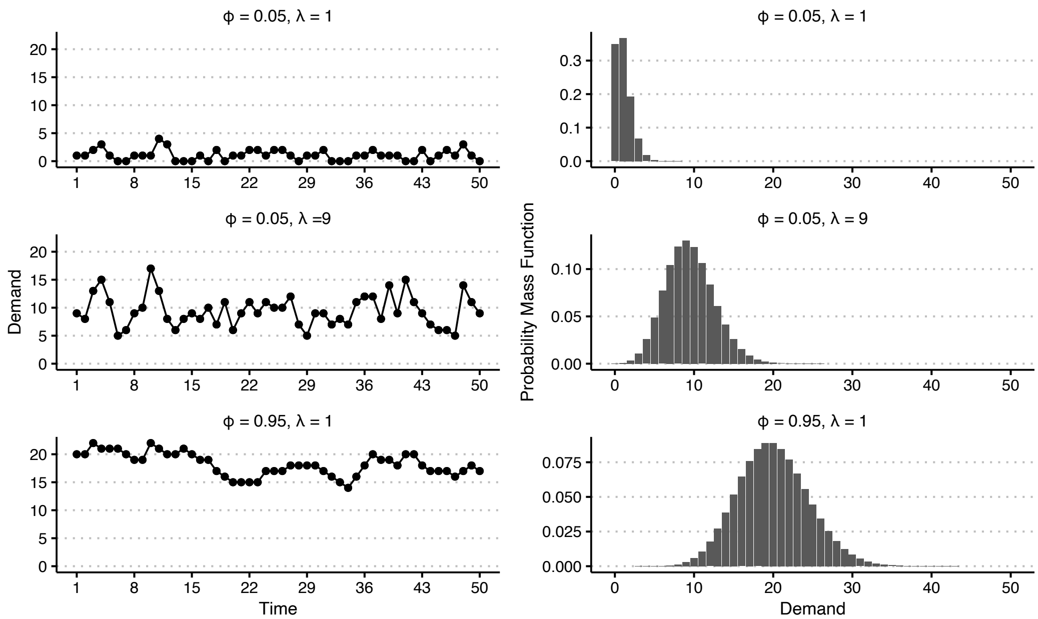

Figure 3: Demand distributions and example time series: Increasing ϕ produces demand processes with trends and higher means. Note: the pmf is produced is calculated from (7). The random seed is identical for all time series.

lead to $SS = F_i^{-1}[b/(b+h)]$ via a standard newsvendor analysis, Boute et al. (2022). Note, here $F_i^{-1}[x]$ is the inverse of the inventory distribution evaluated at x, h is the per unit, per period inventory holding cost and b is the per unit, per period inventory backlog cost. As the inventory distribution is stationary, SS has no influence on the order and inventory variances.

The inventory balance equation is given by

$$i_t = i_{t-1} + q_{t-L} - d_t \tag{11}$$

where i_t is the inventory level at time t, q_{t-L} are the replenishment orders placed in period $t-L$, and $L \geq 1$ is the integer lead-time that includes a sequence of event delay. Note, an order with zero lead time is not deemed to have been received until the next order quantity is determined. Eq. (11) implies the following sequence of events is present: 1) Orders placed in period $t-L$ are received at the start of the period t. 2) Throughout period t, demand is observed and satisfied from inventory. 3) The order-up-to level, s_t, is updated and 4) orders are placed at the end of the period.

The inventory variance calculation requires the determination of: the variance of demand over the lead time, the variance of forecast, and the co-variance between demand and its forecast over the lead time. In the following sections we develop these expressions which depends upon the forecasting method present. We will consider two forecasting methods:

- Real-valued conditional mean forecasts, denoted $\bar{d}$, that originate from an MMSE forecast.
- Integer valued conditional median forecasts, denoted $\tilde{d}$, that minimise the expected absolute forecast error.

In the next section, we define and explore these two forecasting approaches under INAR(1) demand.

3. Forecasting demand over the lead time with the conditional mean

Here we assume the demand forecast minimizes the mean square forecast error over the lead-time and review period, conditional upon d_t. That is, $\hat{d}_{t+k} = \bar{d}_{t+k} = \mathbb{E}[d_{t+k}|d_t]$.

Lemma 2. *Under INAR(1) demand, the conditional mean forecast of demand over the lead time is*

$$\bar{d}_{t,L} = \frac{\phi(1-\phi^L)}{1-\phi} d_t + \frac{L\lambda(1+L)}{2}. \tag{12}$$

Proof. *We start by deriving an exact expression of demand k periods ahead, d_{t+k}. We note*

$$d_{t+k} = \phi \circ d_{t+k-1} + z_{t+k} \text{ and} \tag{13}$$

$$d_{t+k-1} = \phi \circ d_{t+k-2} + z_{t+k-1}. \tag{14}$$

Substituting (14) *into* (13) *recursively (for $d_{t+k-2}, d_{t+k-3}, \ldots$), and collecting ϕ terms* (3) *yields,*

$$d_{t+k} = \phi^k \circ d_t + z_{t+1} + z_{t+2} + \ldots + z_{t+k}. \tag{15}$$

Replacing the future values of z_{t+i} in (15) *with their expectation, $\mathbb{E}[z_{t+i}] = \lambda$, yields,*

$$\bar{d}_{t+k} = \mathbb{E}[\phi^k \circ d_t + z_{t+1} + z_{t+2} + \ldots + z_{t+k} | d_t] = \phi^k d_t + k\lambda. \tag{16}$$

Summing $\sum_{k=1}^{L} \left(\phi^k d_t + k\lambda\right)$ provides (12)*, the stated relation for $\bar{d}_{t,L}$.* □

Note $\bar{d}_{t,L} \in \mathbb{R}$ is increasing in d_t. The first term is increasing in L as $0 \le \phi \le 1$, and independent of λ. The second term simplifies out in subsequent analysis.

3.1. Consequences forecasting INAR(1) demand with the conditional mean

Under i.i.d. integer demand, the demand forecasts based on the conditional mean are constant over time, $\bar{d}_{t,L} = \bar{d}_{t-1,L} = L\mu_d = L\lambda$, and from (8), orders equal demand, $q_t = d_t$. Furthermore, as the variance of the demand is λ when $\phi = 0$, $\mathbb{V}[q_t] = \lambda$ and $\mathbb{V}[i_t] = L\lambda$. As both the demand and the orders are integer processes, so are the inventory levels. However, for correlated INAR(1) demand, the forecast is dynamic ($\bar{d}_{t,L} \neq \bar{d}_{t-1,L}$) and is also real valued ($\bar{d}_{t,L} \in \mathbb{R}$). This means real valued orders and inventory levels are present under correlated demand. These non-integer orders and inventory are incoherent with the integer demand assumption. Hence, we now seek integer forecasts via conditional median forecasts which lead to integer orders and inventory levels.

4. Forecasting INAR(1) demand over the lead time with the conditional median

Let $\tilde{d}_{t+k|t}$ be an integer forecast of the demand k periods ahead, conditional upon d_t. The median of the pmf provides coherent integer forecasts which Freeland and McCabe (2004) claim minimises the expected absolute error conditional upon d_t, $\mathbb{E}[|d_{t+k} - \tilde{d}_{t+k}||d_t]$. The median k periods ahead forecast, $\tilde{d}_{t+k} = X$, where X is the smallest X such that

$$\sum_{x=0}^{X} \mathbb{P}[\tilde{d}_{t+k} = x | d_t] > 1/2. \tag{17}$$

Lemma 3. *The pmf of $\tilde{d}_{t+k} = x$, given d_t is*

$$\mathbb{P}[\tilde{d}_{t+k} = x | d_t] = \binom{d_t}{M_k} e^{\frac{2\lambda\left(\phi^k - 1\right)}{1-\phi}} \left(\phi^k\right)^{M_k} \Gamma(d_t + 1 - M_k) \left(1 - \phi^k\right)^{d_t - M_k} \left(\frac{\lambda\left(\phi^k - 1\right)}{\phi - 1}\right)^{x - M_k}$$
$$\times {}_2\tilde{F}_2 \left[1, -M_k; d_t + 1 - M_k, x + 1 - M_k; \frac{\phi^{-k}\left(\phi^k - 1\right)^2 \lambda}{\phi - 1}\right], \tag{18}$$

where, $x = 0,1,...$ is a non-negative integer, $\Gamma(\cdot)$ is the Gamma function, $M_k = (\tilde{d}_{t+k} \wedge d_t)$ is the minimum of $\tilde{d}_{t+k}$ and d_t, and ${}_2\tilde{F}_2[a_1,a_2;b_1,b_2;z]$ is the regularized generalized hypergeometric function, Mathworld (2021).

Proof. *Freeland and McCabe (2004) and Silva et al. (2009) provide the following expression for the pmf of $\tilde{d}_{t+k}$, given d_t:*

$$\mathbb{P}[\tilde{d}_{t+k} = x | d_t] = e^{\frac{\lambda\left(1-\phi^k\right)}{\phi-1}} \sum_{i=0}^{M_k} \frac{1}{(x-i)!} \binom{d_t}{i} \left(\phi^k\right)^i \left(1-\phi^k\right)^{d_t-i} \left(\frac{\lambda\left(1-\phi^k\right)}{1-\phi}\right)^{x-i}. \tag{19}$$

Algebra allows one to close the sum, producing (18). □

Further analytical work cannot be done with this forecasting technique; however (17) and (18) can be easily implemented and studied numerically in software such as Excel, Mathematica, and R.

4.1. Forecasting i.i.d. integer demand with the conditional median

An i.i.d. integer demand can be modelled within the INAR(1) framework by setting $\phi = 0$. However, the pmf of $\tilde{d}_{t+k}$ when $\phi = 0$ given by (18), is indeterminate and another approach to determine the median forecasts must be taken. When $\phi = 0$, the INAR(1) demand degenerates into an i.i.d. random variable drawn from a Poisson distribution with mean and variance λ those pmf is

$$\mathbb{P}[d_t = x] = \frac{\lambda^x e^{-\lambda}}{x!}. \tag{20}$$

Substituting (20) into (17) produces an implicit expression for the median forecast of the Poisson demand; the median forecasts concur with the smallest X that ensures

$$\sum_{x=0}^{X} \frac{\lambda^x e^{-\lambda}}{x!} = \frac{\Gamma[1+X,\lambda]}{\Gamma[1+X]} > 1/2. \tag{21}$$

Here $\Gamma[\cdot,\cdot]$ is the incomplete Gamma function. Note, when λ is a positive integer, the median of the Poisson demand is equal to its mean λ. When λ is not an integer, (21) implies $\tilde{d}_{t+k} = \lceil\lambda\rceil$ or $\tilde{d}_{t+k} = \lfloor\lambda\rfloor$ depending on the value of λ. Notice, there are no time dependent variables in (21); the median forecast and s_t remains constant over time. The consequences of this ensure the orders always equal the demand under i.i.d. Poisson demand (that is, $q_t = d_t$). This is exactly how the order-up-to policy responds to real valued demand patterns. Further note, $\mathbb{V}[q_t] = \lambda$ and $\mathbb{V}[i_t] = L\lambda$, concurring with the order and inventory variances of the conditional mean forecasted OUT policy.

5. Analysis of the OUT variances with INAR(1) demand

In this section we compare how the OUT policy responds to the INAR(1) demand with the two different forecasting methods. The analysis will focus on the *Bullwhip* and *NSAmp* ratios as they are often used to assess dynamic supply chain performance, Wang and Disney (2016):

$$Bullwhip = \mathbb{V}[q_t]/\mathbb{V}[d_t] \quad \text{and} \quad NSAmp = \mathbb{V}[i_t]/\mathbb{V}[d_t]. \tag{22}$$

When $Bullwhip > 1$ we say there is a bullwhip effect present. The demand variance was given in (6). In §5.1 we determine the order and inventory variance when the conditional mean is used to forecast demand in order to investigate the *Bullwhip* and *NSAmp* measures. §5.2 does the same for the OUT policy with conditional median forecasting.

5.1. Order and inventory variance with conditional mean forecasting

Vassian (1955) shows $\mathbb{V}[i_t]$ is given by the variance of forecast error over lead time:

$$\mathbb{V}[i_t] = \mathbb{V}\left[d_{t,L} - \hat{d}_{t,L}|d_t\right] = \mathbb{V}[d_{t,L}] + \mathbb{V}[\hat{d}_{t,L}|d_t] - 2\text{cov}[d_{t,L}, \hat{d}_{t,L}|d_t]. \tag{23}$$

We need three components: the variance of the demand over the lead time $\mathbb{V}[d_{t,L}]$, the variance of the forecast of the demand over the lead time $\mathbb{V}[\hat{d}_{t,L}]$, and the co-variance between the demand over the lead-time and the forecast of the demand over the lead time $\text{cov}[d_{t,L}, \hat{d}_{t,L}|d_t]$. These are provided in the following three Lemmas.

Lemma 4 (The variance of demand over lead time L). *The variance of the lead-time demand is*

$$\mathbb{V}[d_{t,L}] = L\gamma_0 + 2\sum_{j=1}^{L-1}\sum_{i=1}^{j}\gamma_0\phi^i = \frac{\lambda\left(\left(\phi^2-1\right)L - 2\phi\left(\phi^L - 1\right)\right)}{(\phi-1)^3}. \tag{24}$$

Proof. *The lead time demand is given by*

$$d_{t,L} = \sum_{i=1}^{L} d_{t+i} = d_{t+1} + d_{t+2} + \ldots + d_{t+L}. \tag{25}$$

The variance of demand over the lead time is calculated from

$$\begin{aligned}\mathbb{V}[d_{t,L}] = \mathbb{V}[d_{t+1} + d_{t+2} + \ldots + d_{t+L}] = \sum_{i=1}^{L}(\mathbb{V}[d_{t+i}] + 2cov[d_{t+1}, d_{t+2}] + \ldots + 2cov[d_{t+1}, d_{t+L}] \\ + 2cov[d_{t+2}, d_{t+3}] + \ldots + 2cov[d_{t+2}, d_{t+L}) + \ldots + 2cov[d_{t+L-1}, d_{t+L}]).\end{aligned} \tag{26}$$

Using (6) *in* (26) *yields the variance of the demand over the lead time,*

$$\mathbb{V}[d_{t,L}] = L\gamma_0 + 2(\gamma_1 + \gamma_2 + \ldots\gamma_{L-1}) + 2(\gamma_1 + \gamma_2 + \ldots\gamma_{L-2}) + \ldots + 2\gamma_1 \tag{27}$$

Using (6) *and the telescoping method, the nested sum in* (27) *becomes the stated relation* (24). □

Lemma 5 (Variance of the forecast over the lead time). *First, note the forecast $\hat{d}_{t,L} = \bar{d}_{t,L}$ with conditional mean forecasting. From* (6) *and* (12)*, the variance of the forecast over the lead time is*

$$\mathbb{V}\left[\hat{d}_{t,L}|d_t\right] = \mathbb{V}\left[d_t\frac{\phi(1-\phi^L)}{1-\phi} + \frac{L\lambda(1+L)}{2}\right] = \mathbb{V}\left[d_t\frac{\phi(1-\phi^L)}{1-\phi}\right] = \frac{\lambda}{1-\phi}\left(\frac{\phi(1-\phi^L)}{1-\phi}\right)^2. \square \tag{28}$$

Eq. (28) is increasing in L as $0 \leq \phi \leq 1$.

Lemma 6 (Covariance of the demand and its forecast over L). *The covariance of demand over lead time and its forecast is calculated as,*

$$cov\left[d_{t,L}, \hat{d}_{t,L}|d_t\right] = cov\left[d_{t+1} + d_{t+2} + \ldots + d_{t+L}, \hat{d}_{t,L}|d_t\right]. \tag{29}$$

By substituting (12) *into* (29)*, using* (6) *and the additive law of covariance (the covariance of a random variable with a sum of random variables is the sum of the covariances with each of the random variables) we have:*

$$\begin{aligned} cov\left[d_{t,L}, \hat{d}_{t,L} | d_t\right] &= cov\left[d_{t+1} + ... + d_{t+L}, d_t \frac{\phi(1-\phi^L)}{1-\phi} + \frac{L\lambda(1+L)}{2}\right] \\ &= \frac{\phi(1-\phi^L)}{1-\phi}(\gamma_1 + \gamma_2 + ... + \gamma_L) = \gamma_0 \left(\frac{\phi\left(1-\phi^L\right)}{1-\phi}\right)^2. \square \end{aligned} \tag{30}$$

Finally, we can now provide the inventory variance expression in Proposition 1:

Proposition 1 (Variance of the inventory levels). *The inventory variance is given by*

$$\mathbb{V}[i_t] = \frac{\lambda}{1-\phi}\left(L + 2\phi\left(\frac{\phi^L + L(1-\phi) - 1}{(\phi-1)^2}\right) - \left(\frac{\phi(1-\phi^L)}{1-\phi}\right)^2\right). \tag{31}$$

Proof. *Eq.* (23) *highlighted the variance of the inventory levels is given by the variance of the forecast error over the lead time. Substituting* (6)*,* (24)*,* (28)*, and* (30) *into* (23) *yields* (31)*.* □

The *NSAmp* measure, plotted in Figure 4, for the OUT policy under INAR(1) demand is the same as the *NSAmp* measure under AR(1) demand, Disney and Lambrecht (2008). Notice, as $\phi \to 1$, the demand variance $\mathbb{V}[d_t] \to \infty$. Together with a finite inventory variance, this means $NSAmp \to 0$ as $\phi \to 1$. Further $NSAmp \to L$ as $\phi \to 0$.

Proposition 2 (Variance of the orders). *The variance of the replenishment orders is given by*

$$\mathbb{V}[q_t] = \frac{\lambda}{1-\phi}\left(1 + 2\phi\left(1-\phi^L\right)\left(1 + \frac{\phi\left(1-\phi^L\right)}{1-\phi}\right)\right). \tag{32}$$

Proof. *First note, substituting s_t and s_{t-1} from* (9) *into* (8) *yields*

$$q_t = \hat{d}_{t,L} - \hat{d}_{t-1,L} + d_t. \tag{33}$$

Using $\hat{d}_{t,L}$ and $\hat{d}_{t-1,L}$ from (12) *in* (33) *and collecting together like terms we obtain:*

$$\begin{aligned} q_t =& \frac{L\lambda}{1-\phi} + \phi\left(d_t - \frac{\lambda}{1-\phi}\right)\left(\frac{1-\phi^L}{1-\phi}\right) - \frac{L\lambda}{1-\phi} - \phi\left(d_{t-1} - \frac{\lambda}{1-\phi}\right)\left(\frac{1-\phi^L}{1-\phi}\right) + d_t \\ =& (d_t - d_{t-1})\phi\left(\frac{1-\phi^L}{1-\phi}\right) + d_t. \end{aligned} \tag{34}$$

The variance of order quantity is calculated as follows:

$$\begin{aligned} \mathbb{V}[q_t] =& \mathbb{V}\left[\frac{\phi d_t(1-\phi^L)}{1-\phi} - \frac{\phi d_{t-1}(1-\phi^L)}{1-\phi} + d_t\right] \\ =& \mathbb{V}\left[\frac{\phi d_t(1-\phi^L)}{1-\phi}\right] + \mathbb{V}\left[\frac{\phi d_{t-1}(1-\phi^L)}{1-\phi}\right] + \mathbb{V}[d_t] + 2cov\left[\frac{\phi d_t(1-\phi^L)}{1-\phi}, d_t\right] \\ & - 2cov\left[\frac{\phi d_t(1-\phi^L)}{1-\phi}, \frac{\phi d_{t-1}(1-\phi^L)}{1-\phi}\right] - 2cov\left[\frac{\phi d_{t-1}(1-\phi^L)}{1-\phi}, d_t\right]. \end{aligned} \tag{35}$$

As $\mathbb{V}[d_{t-k}] = \gamma_0$ and $\forall k \geq 1$, $cov[d_t, d_{t-k}] = \gamma_k$, (35) *reduces to* (32)*.* □

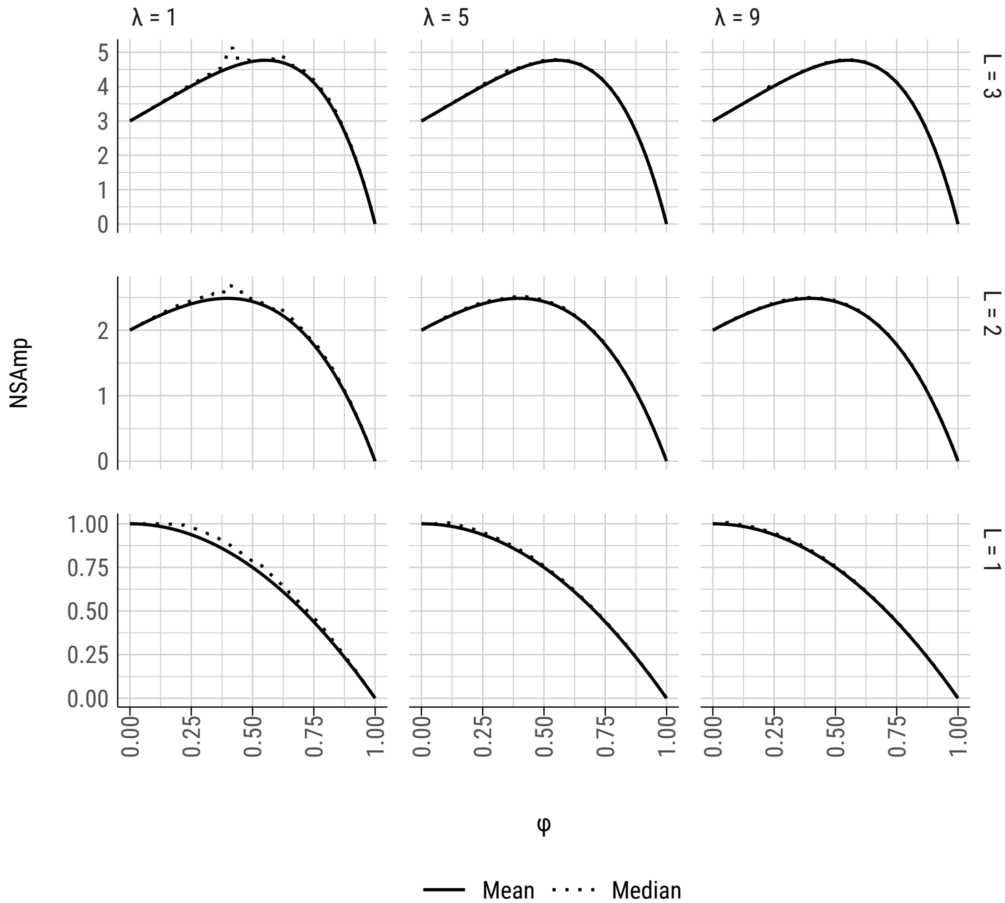

Figure 4: The *NSAmp* maintained by the OUT policy under INAR(1) demand with conditional mean and conditional median forecasting.

By substituting (6) and (32) into (33) we obtain the following expression for the bullwhip effect,

$$Bullwhip = 1 + 2\phi\left(1-\phi^L\right)\left(1+\frac{\phi(1-\phi^L)}{1-\phi}\right). \tag{36}$$

Eq. (36) is interesting. First the Poisson distribution, λ, has no influence on the bullwhip effect. Second, (36) has the same structural form as the *Bullwhip* generated by the OUT policy with MMSE forecasting under AR(1) demand, Disney and Lambrecht (2008). This should not be a surprise; the autocorrelation function (ACF) of the demand process ACF=ϕ^k, given in (6), has exactly the same form as the ACF for the real valued AR(1) process. Furthermore, as $0 \le \phi \le 1$, (36) is increasing in L and always $Bullwhip > 1$ indicating bullwhip always exists under an INAR(1) demand process with conditional mean forecasts, regardless of ϕ and L.

Figure 5 illustrates the impact of the demand process parameter and lead time on the bullwhip effect confirming it is always present and increases in the lead time L. Luong (2007) provides an upper bound for the bullwhip effect, valid for all lead times and real-valued demand and forecasts, given by $(1+\phi)/(1-\phi)$ that is also relevant here. This upper bound is also plotted in Figure 5.

5.2. Order and inventory variance with conditional median forecasting

As no further analytical work on the conditional median forecasts is possible, we resort to a numerical investigation of the orders and inventory variance. We built a simulation in the R

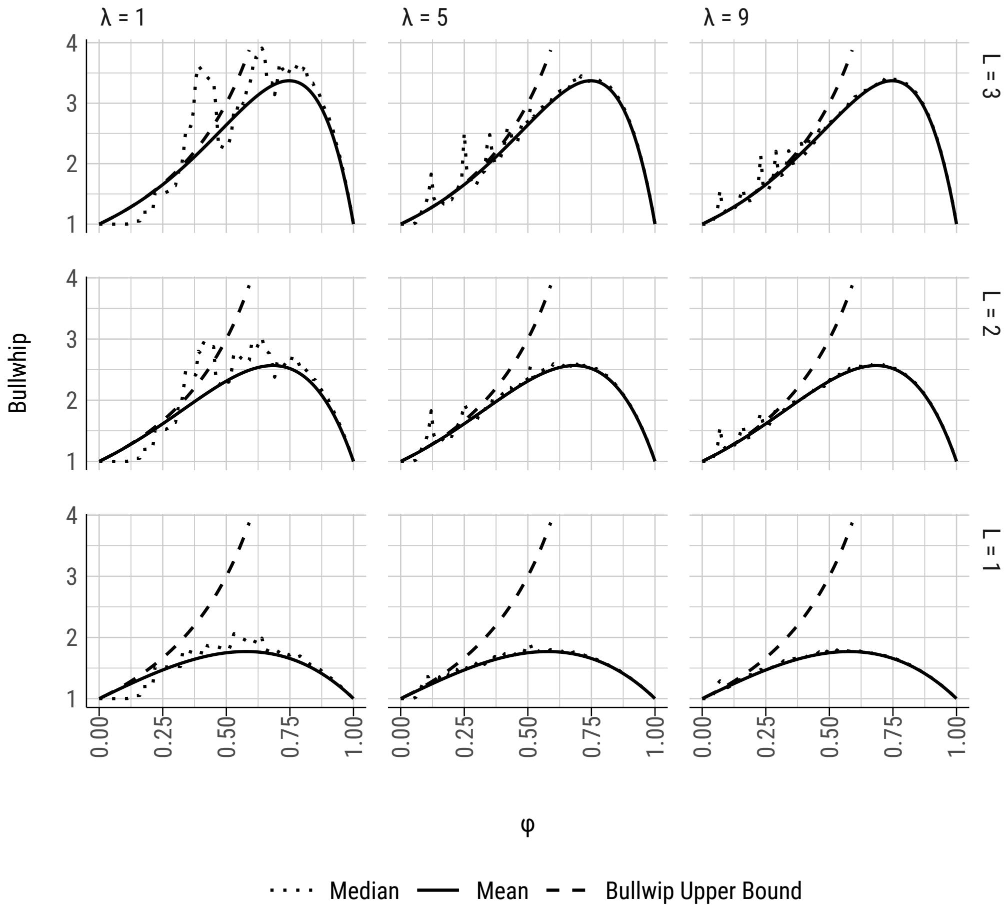

Figure 5: The bullwhip effect in the OUT policy under INAR(1) demand.

software that ran on a HAWK High-Performance Computing Cluster. To generate the demands in each period t, we first generated randomly Poisson distributed error terms, z_t and constructed demand series using (1). Then, k period ahead conditional median forecasts were generated using (17); these were summed to create integer forecasts of demand over the lead time. Following that, inventory balance and orders are calculated using (11) and (8), respectively. Finally, the *Bullwhip* and *NSAmp* ratios are calculated using (22). The parameter values used in the simulation are $0 < \lambda = [1,9] < 1$ and $0 < \phi < 1$ in steps of 0.01. For all combinations of λ and ϕ, a time series of one million observations is generated. We have plotted the results of this exercise in Figure (4) (for the *NSAmp* ratio) and Figure (5) (for the *Bullwhip* ratio).

Figure 4 revealed the *NSAmp* expressions under conditional mean forecasting is a good predictor of the *NSAmp* measure when conditional median forecasting is present, especially with high volume integer demand (i.e. when $\lambda >> 1$). As the conditional mean forecast minimises the mean squared error forecasts over the lead-time and review period, the conditional mean *NSAmp* represents a lower bound of conditional median *NSAmp*; lower demand volumes (small λ and ϕ near 0.5) increase the conditional median *NSAmp*.

The conditional median *Bullwhip* curves in Figure 5 were somewhat more complex. When ϕ is near zero, the *Bullwhip* curves remain unchanged from the i.i.d. case, close to unity, probably as a result of the forecast remaining unchanged from the i.i.d. case. In the region $\phi \approx 0.1$ to $\phi \approx 0.8$ there is a period of seemingly erratic *Bullwhip* which may be either above or below the conditional mean *Bullwhip* curve; it may also be above the AR(1) bullwhip upper bound, Luong (2007). When

ϕ is close to unity, above $\phi \approx 0.8$, the conditional median *Bullwhip* curves closely resembles the conditional mean *Bullwhip* curve. We do not know whether a conditional median ever produces $Bullwhip < 1$.

6. Economic performance of the OUT policy under i.i.d. integer demands

When $\phi = 0$, we have i.i.d. Poisson distributed demands, the demand forecasts are a constant and it is easy to verify the inventory pmf is a simple reflection and translation of the pmf of the sum of demand over the lead-time:

$$\mathbb{P}[i_t = x] = \frac{(L\lambda)^{L\lambda+\bar{i}-x} e^{-L\lambda}}{(L\lambda+\bar{i}-x)!}. \tag{37}$$

We can determine the mean and variance of the inventory levels directly from the mean and variance of the sum of L Poisson distributed random variables:

$$\mathbb{E}[i_t] = \sum_{x=-\infty}^{L\lambda+\bar{i}} \frac{(L\lambda)^{L\lambda+\bar{i}-x} e^{-L\lambda}}{(L\lambda+\bar{i}-x)!} = \bar{i}, \text{ and } \mathbb{V}[i_t] = \sum_{x=-\infty}^{L\lambda+\bar{i}} \frac{(L\lambda)^{L\lambda+\bar{i}-x} e^{-L\lambda}}{(L\lambda+\bar{i}-x)!}(x-\mathbb{E}[i_t])^2 = L\lambda. \tag{38}$$

The expected per period inventory holding and backlog costs (see (10)) can be then obtained from (37) as follows:

$$\begin{aligned}\mathbb{E}[C_t^i] &= \sum_{x=-\infty}^{0} \frac{b(-x)\left(e^{L\lambda}(L\lambda)^{\bar{i}+L\lambda-x}\right)}{(\bar{i}+L\lambda-x)!} + \sum_{x=1}^{L\lambda+\bar{i}-1} \frac{hx\left(e^{-L\lambda}(L\lambda)^{\bar{i}+L\lambda-x}\right)}{(\bar{i}+L\lambda-x)!} + he^{-L\lambda}(\bar{i}r+L\lambda) \\ &= \frac{(b+h)e^{-L\lambda}\left((L\lambda)^{\bar{i}+L\lambda+1} + \bar{i}e^{L\lambda}\Gamma[\bar{i}+L\lambda+1, L\lambda]\right)}{\Gamma[\bar{i}+L\lambda+1]} - b\bar{i}\end{aligned} \tag{39}$$

Using (39) it is easy to conduct a search on the integers for the optimal $i^\star$, the $\bar{i}$ that minimises (39). Conceptually, $i^\star$ is the safety stock that minimises the expected per period inventory costs as for most reasonable cost settings will imply $0 \leq i^\star \leq L\lambda$. This search is facilitated by the fact that the backlog costs are non-increasing in $\bar{i}$ and the holding costs are non-decreasing in $\bar{i}$. This means there is only one minimum (or two consecutive minimums) in the inventory costs. The search results for $i^\star$ are shown in Figure 6 (left panel) for integer L and integer λ. The safety safety stock $i^\star$ appears to be symmetrical about $L = \lambda$ and increasing in both λ and L.

In a like manner, due to the constant forecasts under i.i.d. INAR(1) demand, (8) shows that the order pmf is simply equal to the demand pmf and and $Bullwhip = 1$. The per period production cost C_t^q, with a nominal hours <u>u</u>nit cost of u within a nominal capacity of K and flexible per unit overtime cost of um, where m is the overtime <u>m</u>ultiplier, (Boute et al. (2022)):

$$C_t^q = uK + um[q_t - K]^+. \tag{40}$$

Using (20), the expected per period capacity costs are given by

$$\mathbb{E}[C_t^q] = uK + \sum_{x=k}^{\infty} \frac{mu(x-K)e^{-\lambda}\lambda^x}{x!} \tag{41}$$

$$= u\left(\frac{e^{-\lambda}\lambda m\left(\lambda^k - e^{\lambda}\Gamma[k+1,\lambda]\right)}{\Gamma[k+1]} + \frac{m\Gamma[k+1,\lambda]}{\Gamma[k]} + K + m(\lambda - K)\right) \tag{42}$$

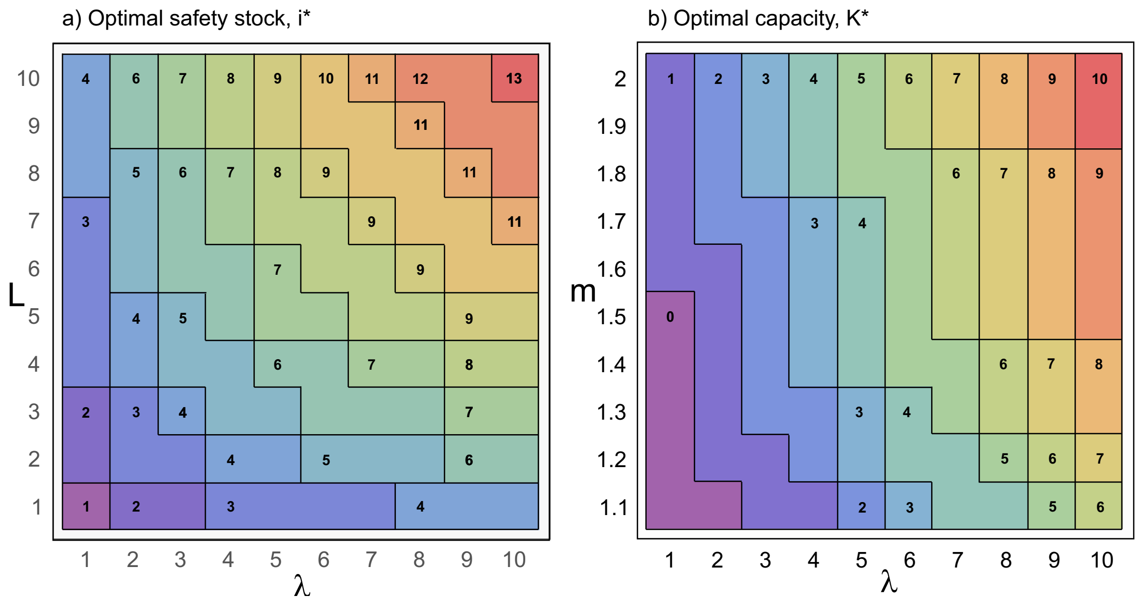

Figure 6: Optimal OUT policy settings. Panel a) The safety stock, $i^{\star}$, required to minimise the inventory costs when $h=1, b=9$ for different demand λ and lead times L. Panel b) The optimal capacity, $K^{\star}$, required to minimise the capacity costs costs when $u=4$ for different demand λ and over-time multiplier m.

Notice, in (41), the first addend uK is increasing in K, and the second addend, the sum, is decreasing in K. Thus, there is a unique K (or at most two consecutive Ks) that minimise the expected capacity costs. Consequently, (41) can also be minimised effectively via a line search on the integers for K. Doing so, yields Figure 6 (right panel) which shows the $K^{\star}$, the K that minimises (41). $K^{\star}$, is both increasing in λ and m.

7. Conclusions

We have examined the *Bullwhip* and *NSAmp* behavior of the OUT policy under an integer-valued INAR(1) demand series with two different forecasts methodologies; one based on the conditional mean, the other forecast on the conditional median. The variance ratios were derived for conditional mean forecasts under integer demand and found to be the same as for the corresponding real-valued demand. This should be expected as the results for the real demand variance are *distribution free*. The Poisson distribution parameter λ had no impact on the variance ratios under INAR(1) demand. However, the real order and inventory levels produced are incoherent with the integer demand. The bullwhip effect always exists regardless of the auto-regressive parameter and the lead time. There exists a lower bound, $Bullwhip > 1$, and an upper bound which is a function of the auto-regressive parameter, ϕ. For a given value of ϕ, the upper bound represents the maximum value of the bullwhip effect regardless of the lead time L. The upper bound is tight when ϕ is small.

When conditional median forecasts are used, bullwhip is somewhat more erratic and can deviate (positively and negatively) significantly from both the conditional mean bullwhip and the upper bound. The *Bullwhip* and *NSAmp* expressions can be used with confidence in high volume settings. However, low values of ϕ and λ lead to intermittent demand series that contain a high proportion of zeros and to time series where the integer effects become more significant. Bullwhip seems to always exists for INAR(1) demand; the inventory variance (and the *NSAmp* measure) for conditional mean forecasts is a lower bound for the inventory variance under conditional median forecasts. When the demand is i.i.d., a constant forecast is produced by both forecasting methods that meant $q_t = d_t$. For this case we were able to extend our variance ratio analysis to include an economic study of the system that included a search for the target safety stock to minimised the expected inventory holding and backlog costs and the target capacity level to minimise the regular labour and overtime costs.

Acknowledgements

This research was undertaken using the supercomputing facilities at Cardiff University operated by Advanced Research Computing at Cardiff (ARCCA) on behalf of the Cardiff Supercomputing Facility and the HPC Wales and Supercomputing Wales (SCW) projects. We acknowledge the support of the latter, which is part-funded by the European Regional Development Fund (ERDF) via the Welsh Government.

Appendix A. Auto-covariance function of the demand

We can show that the mean of d_t is given by,

$$\mathbb{E}[d_t] = \mathbb{E}[\phi \circ d_{t-1} + z_t] = \mathbb{E}[\phi \circ d_{t-1}] + \mathbb{E}[z_t] = \phi \mathbb{E}[d_{t-1}] + \lambda. \tag{A.1}$$

INAR(1) processes are stationary, $\mathbb{E}[d_t] = \mathbb{E}[d_{t-1}]$ and (A.1) reduces to $\mathbb{E}[d_t] = \lambda/(1-\phi)$. As the variance of a sum is the the sum of the variances of the addends and twice the covariance between the addends, the variance of demand at period t is

$$\mathbb{V}[d_t] = \mathbb{V}[\phi \circ d_{t-1} + z_t) = \mathbb{V}[\phi \circ d_{t-1}] + \mathbb{V}[z_t] + 2\text{cov}[d_{t-1}, z_t]. \tag{A.2}$$

Algebra then leads to

$$\begin{aligned}
\mathbb{V}[d_t] &= \mathbb{V}[\phi \circ d_{t-1}] + \lambda && (\text{As cov}[d_{t-1}, z_t] = 0) \\
&= \phi \mathbb{V}[d_{t-1}] + \lambda && (\mathbb{V}[\phi \circ d_{t-1}] = \phi \mathbb{V}[d_{t-1}]) \\
&= \phi \mathbb{V}[d_t] + \lambda && (\text{As demand is stationary, } \mathbb{V}[d_{t-1}] = \mathbb{V}[d_t]) \\
&= \lambda/(1-\phi). && (\text{After collecting together terms})
\end{aligned}$$

By recursive substitutions of d_{t-k} for $k \geq 1$, (1) can be written as

$$d_t = \phi^k \circ d_{t-k} + \sum_{j=0}^{k-1} \phi^j z_{t-j}. \tag{A.3}$$

The auto-covariance of lag $k \geq 1$ can then be calculated as

$$\begin{aligned}
\gamma_k = \text{cov}[d_t, d_{t-k}] &= \text{cov}\left[\phi^k \circ d_{t-k} + \sum_{j=0}^{k-1} \phi^j \circ z_{t-j}, d_{t-k}\right] \\
&= \phi^k \text{cov}[d_{t-k}, d_{t-k}] + \text{cov}\left[\sum_{j=0}^{k-1} \phi^j \circ z_{t-j}, d_{t-k}\right].
\end{aligned} \tag{A.4}$$

As the correlation between d_{t-k} and z_{t-j} for all $j \leq k-1$ is equal to zero, the covariance term $\text{cov}\left[d_{t-k}, \sum_{j=0}^{k-1} \phi^j \circ z_{t-j}\right] = 0$. The auto-covariance function of lag $k \geq 1$ for INAR(1) demand is

$$\gamma_k = \phi^k \text{cov}[d_{t-k}, d_{t-k}] = \phi^k \gamma_0. \tag{A.5}$$

References

Al-Osh, M., Alzaid, A.A., 1988. Integer-valued moving average (INMA) process. Statistical Papers 29, 281–300.

Boute, R., Disney, S.M., Gijsbrechts, J., Mieghem, J.A.V., 2022. Dual sourcing and smoothing under non-stationary demand time series: Re-shoring with SpeedFactories. Forthcoming in Management Science doi: 10.1287/mnsc.2020.3951.

Box, G.E., Jenkins, G.M., Reinsel, G.C., Ljung, G.M., 2015. Time Series Analysis: Forecasting and Control. John Wiley & Sons.

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

Chen, F., Ryan, J.K., Simchi-Levi, D., 2000b. The impact of exponential smoothing forecasts on the bullwhip effect. Naval Research Logistics 47, 269–286.

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

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

Duc, T.T.H., Luong, H.T., Kim, Y.D., 2008. A measure of bullwhip effect in supply chains with a mixed autoregressive-moving average demand process. European Journal of Operational Research 187, 243–256.

Freeland, R., McCabe, B., 2004. Forecasting discrete valued low count time series. International Journal of Forecasting 20, 427–434.

Graves, S.C., 1999. A single-item inventory model for a non-stationary demand process. Manufacturing & Service Operations Management 1, 50–61.

Hosoda, T., Disney, S.M., 2006. The governing dynamics of supply chains: The impact of altruistic behaviour. Automatica 42, 1301–1309.

Lee, H., Padmanabhan, P., Whang, S., 1997. Information distortion in a supply chain: the bullwhip effect. Management Science 43, 543–558.

Lee, H., So, K., Tang, C., 2000. The value of information sharing in a two-level supply chain. Management Science 46, 626–643.

Luong, H., 2007. Measure of bullwhip effect in supply chains with auto-regressive demand process. European Journal of Operational Research 180, 1086–1097.

Mathworld, 2021. Regularized hypergeometric function. https://mathworld.wolfram.com/RegularizedHypergeometricFunction.html.

Ristić, M.M., Nastić, A.S., 2012. A mixed INAR(p) model. Journal of Time Series Analysis 33, 903–915.

Silva, M.E.D., Oliveira, V.L., 2004. Difference equations for the higher-order moments and cumulants of the INAR(1) model. Journal of Time Series Analysis 25, 317–333.

Silva, N., Pereira, I., Silva, E.M., 2009. Forecasting in INAR(1) model. Revstat – Statistical Journal 7, 119–134.

Vassian, H.J., 1955. Application of discrete variable servo theory to inventory control. Journal of the Operations Research Society of America 3, 272–282.

Wang, X., Disney, S.M., 2016. The bullwhip effect: Progress, trends and directions. European Journal of Operational Research 250, 691–701.

Bahman Rostami-Tabar is a Senior Lecturer in Management Science at Cardiff Business School, Cardiff University. Bahman's research goals are directed toward the use of Operations Research techniques to improve decision making in supply chains and operations management, in order to positively contribute to both industrial and societal advancements.

Stephen Disney is Professor of Operations Management and Head of the Science, Innovation, Technology and Entrepreneurship Department at the University of Exeter Business School. His research interests involve the application of control theory and statistical techniques to operations management and supply chain scenarios to investigate their dynamic, stochastic, and economic performance. Stephen has a particular interest in the bullwhip effect, forecasting, inventory management and closed loop supply chains.