Li, Q., Disney, S.M. and Gaalman, G., (2014), "Avoiding the bullwhip effect using Damped Trend forecasting and the Order-Up-To replenishment policy", International Journal of Production Economics, Vol. 149, pp3–16. DOI: 10.1016/j.ijpe.2013.11.010.

Avoiding the bullwhip effect using Damped Trend forecasting and the Order-Up-To replenishment policy

Qinyun Li [a], Stephen M. Disney [a], Gerard Gaalman [b]

a) Cardiff Business School, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK.
b) Department of Operations, Faculty of Economics and Business, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands.

Abstract

We study the Damped Trend forecasting method and its bullwhip generating behaviour when used within the Order-Up-To (OUT) replenishment policy. Using *z*-transform transfer functions we determine complete stability criteria for the Damped Trend forecasting method. We show that this forecasting mechanism is stable for a much larger proportion of the parametrical space than is generally acknowledged in the literature. We provide a new proof to the known fact that the Naïve, Exponential Smoothing and Holts Method forecasting, when used inside the OUT policy, will always generate bullwhip for every possible demand process, for any lead-time. Further, we demonstrate the Damped Trend OUT system behaves differently. Sometimes it will generate bullwhip and sometimes it will not. Bullwhip avoidance occurs when demand is dominated by low frequency harmonics in some instances. In other instances bullwhip avoidance happens when demand is dominated by high frequency harmonics. We derive sufficient conditions for when bullwhip will definitely be generated and necessary conditions for when bullwhip may be avoided. We verify our analytical findings with a numerical investigation.

Keywords: Damped Trend Forecasting, Order-Up-To Replenishment Policy, Bullwhip, Stability, *z*-transform, Fast Fourier Transform, Frequency Response.

1. Introduction

The Damped Trend (DT) forecasting method generalises the Holts model for forecasting a linear trend by adding a damping parameter ϕ to the trend component. Although such exponential smoothing systems with damping parameters had been noted earlier (see for instance, Gardner (1985), Gilchrist (1976), and Roberts (1982)), Gardner and McKenzie (1985) were the first to present both a theoretical and an empirical investigation of the system. Since then, it has often been promoted as the most accurate forecasting technique in the so-called M-competitions (Makridakis and Hibon, 2000). Gardner and McKenzie (2011) find the DT method is the best method for 84% of the 3003 time series in the M3 forecasting competition when using local initial values. It was the best method 70% of the time when using global initial values. Fildes et al. (2008) concluded that DT forecasting can "reasonably claim to be a benchmark forecasting method for all others to beat". The great virtual of DT is that future forecasts are not simply flat line extensions of the current, next period forecast. It is able to detect and forecast both linear and exponential trends. The DT forecasting methodology also contains at least eleven different forecasting methods when all of the three parameters are selected from the real $[0,1]$ interval (Gardner and McKenzie, 2011). This makes it a powerful and very general forecasting approach for short term demand data as tuning the DT parameters effectively automates model selection. In this paper, as they are industrially popular, we are particularly interested in studying the performance of Naïve, Simple Exponential Smoothing, Holts Method and Damped Trend forecasting procedures.

A series of papers (Gardner, 1985, 1990; Gardner and McKenzie, 1985, 1989) have proposed restricting the damping parameter to the $[0,1]$ interval, and that the rate of decay $(1-\phi)$ increases with the noise in the series because the difference between a damped and a linear trend can be substantial over long time horizons. When $\phi>1$, both Gardner (1985) and Gardner and McKenzie (1985) explain that the forecast exhibits an exponential growth over time and is probably a dangerous option to use in an automatic forecasting procedure. However, both Tashman and Kruk (1996) and Taylor (2003) argue that there can be value in allowing $\phi>1$ as it could be suited to time series with strong increasing trends. Roberts (1982) mentions that a negative ϕ is possible, but generally negative damping parameters are not discussed in the literature. In this paper, we will investigate the general DT method without any restrictions on the range of values any parameter can take.

Although DT method is claimed to be superior to many other methods from a forecasting perspective, only a few studies are available that demonstrate the managerial importance of DT forecasting in the supply chain, inventory or operations management fields. Snyder et al. (2002) studies the exponential smoothing family of forecasting methods (including Damped Trend) in the Order-Up-To inventory control policy. They use a bootstrap method to determine the total lead-time demand distribution and measure performance via the fill rate. Acar and Gardner (2012) investigate the use of the DT method in a real supply chain based on the trade-offs recommended by Gardner (1990) and show that the DT method outperforms Simple Exponential Smoothing and the Holts method.

We study forecasting techniques from a supply chain perspective, using the forecasts inside the Order-Up-To (OUT) replenishment policy and evaluating performance via the bullwhip criterion. This focus allows us to sharpen the results of Dejonckheere et al. (2003) for the Naïve and the Simple Exponential Smoothing forecasting techniques and to introduce new results for the Holts and Damped Trend forecasting methods.

Our specific focus is to evaluate the use of the DT forecast for use within the linear OUT replenishment policy (Chen et al., 2000). This industrially popular policy has been selected for our study as it is the optimal linear policy for minimising local inventory costs and it operates on the same discrete time periodic basis as the DT forecasting method. The OUT policy requires two forecasts, one forecast of demand over the lead-time, and one forecast of demand in the period after the lead-time. The OUT policy is a popular rule as, for any set of forecasts the policy determines replenishment quantities that minimise the variance of the inventory levels, Vassian (1955) and Hosoda and Disney (2006).

Note also that the OUT policy is not restricted to certain demand patterns. This suggests the OUT policy would be suitable when demand contains significant trends. The minimised inventory variance leads to reduced safety stock requirements in supply chains (and lower inventory costs) in order to meet availability targets. However, most OUT policy settings result in a phenomenon where the variance of the orders is greater than the variance of the demand, a.k.a. the bullwhip effect. This bullwhip effect is costly in supply chains as it creates costs either in the form of labour idling and over-time or excessive and unused capacity. It has been shown that the bullwhip effect is related to the variance of forecasts as well as the variance of the forecast errors of demand over the lead-time and review period, Chen et al. (2000), Dejonckheere et al. (2003).

In this paper, since the combined DT / OUT system consists of six linear discrete time difference equations, the *z*-transform and Fourier transform approaches are used for the analysis. *z*-transform approaches have a long history in production control and inventory management. Vassian (1955) appears to be the first to apply the *z*-transform to an inventory control problem. Brown (1963) is perhaps the first textbook that details how to apply *z*-transform techniques to forecasting problems. Adelson (1966) studies the dynamic behaviour of coupled forecasting and scheduling systems. More recently, Popplewell and Bonney (1987) study the structure of MRP systems with *z*-transforms and Hoberg, Thonemann and Bradley (2007) use *z*-transforms to study bullwhip behaviour in linear supply chains.

The use of the Fourier Transform for studying the dynamic behaviour of forecasting and replenishment systems is less common: Dejonckheere et al. (2003) use the Fourier transform to investigate the behaviour of the OUT policy with exponential smoothing and moving average forecasts. Ouyang and Daganzo (2006) characterize the bullwhip effect in linear supply chains with Robust Control techniques. The frequency response approach is particularly potent as it is able to generate results that are applicable for any demand process. This is because all demand processes can be decomposed into a set of harmonic frequencies via the Fourier Transform. By understanding how the system reacts to the complete set of harmonic frequencies (via the Amplitude Ratio within the Frequency Response Plot, a.k.a. the Bode Plot) we are able to gain insights that are valid for all possible demand patterns. Many of the results that we obtain are also valid for any lead-time.

Section 2 briefly introduces the DT forecasting mechanism and derives its discrete time transfer function. We then study the stability boundaries of DT forecasting mechanism in Section 3. In Section 4, we incorporate the DT forecasting methodology into the OUT replenishment policy, and develop a discrete-time *z*-transform transfer function representation of the combined forecasting and replenishment system. Then, in Section 5, we analyse the frequency response plot in order to ascertain the bullwhip performance of the system. Section 6 provides numerical simulation results confirming our theoretical findings. Section 7 concludes.

2. The Damped Trend forecasting method

There appears to be no commonly accepted set of notation for the Damped Trend forecasting method. Different authors use different notation and several different formulations of the DT model exist. We prefer to use the recursion form due to Gardner and McKenzie (1985) shown in (1) but have made slight changes to notation in order to avoid $\{S, T, L\}$ which already have implied meanings in the supply chain management literature. This three dimensional form of the DT model is the most general form in the literature.

$$\left.\begin{aligned} \hat{a}_t &= (1-\alpha)\left(\hat{a}_{t-1}+\phi\hat{b}_{t-1}\right)+\alpha d_t \\ \hat{b}_t &= (1-\beta)\phi\hat{b}_{t-1}+\beta\left(\hat{a}_t-\hat{a}_{t-1}\right) \\ \hat{d}_{t,t+k} &= \hat{a}_t+\hat{b}_t\gamma(k,\phi) \end{aligned}\right\}. \tag{1}$$

In the linear discrete time difference equations of (1), d_t is the time series being forecasted, in our context we refer to it as "demand at time *t*". $\hat{a}_t$ is the current estimate of the level, exponentially smoothed by the constant level α . $\hat{b}_t$ is the current estimate of the trend, exponentially smoothed by the constant β . $\{\alpha,\beta\}$ are “smoothing parameters”. ϕ is the damping parameter that can be interpreted as a measure of the persistence of the trend. k is the number of periods ahead that the forecast is required to predict. $\hat{d}_{t,t+k}$ is the forecast, made at time *t*, of the demand in the period $t+k$ and $\gamma(k,\phi)=\sum_{i=1}^{k}\phi^i=\phi\left(\phi^k-1\right)/\left(\phi-1\right)$. The first equation of (1) suggests that the estimated demand consists of a time dependant “level” component, the second equations tracks a trend component and the third equation combines the level and trend estimates to make a forecast *k* periods ahead. The behaviour of $\gamma\left(k,\phi\right)$ is quite rich. When $\phi>1$ then $\gamma\left(k,\phi\right)$ exhibits positive exponential growth over *k*. $\phi=1$ implies a $\gamma\left(k,\phi\right)$ with positive linear growth over *k*. $0<\phi<1$ produces a positive damped exponential growth that approaches $\phi/\left(1-\phi\right)$ as $k\to\infty$. When $\phi=0$ then $\gamma\left(k,\phi\right)=0$. When $-1<\phi<0$, $\gamma\left(k,\phi\right)$ has a two period oscillation that is always negative and converges to $\phi/\left(1-\phi\right)$. $\phi<-1$ results in a $\gamma\left(k,\phi\right)$ with a two period oscillation that alternates between positive and negative numbers with ever increasing amplitude over *k*.

Several well-known forecasting approaches are encapsulated within the DT model (Gardner and McKenzie, 2011). These include the Holts method when $\phi=1$ where there is no damping of the trend component, Simple Exponential Smoothing (SES) when $\beta=0$ and $\phi=0$, and Naïve forecasting when $\alpha=1$, $\beta=0$ and $\phi=0$, see Table 1. We study the system performance when the OUT policy exploits either DT forecasting, or these special cases, to predict demand over the lead-time and review period.

Forecasting method	Parameter settings	Notes
Holts method	$\phi=1$	Setting $\phi=1$, $\hat{d}_{t,t+k}=\hat{a}_t+k\hat{b}_t$. Future forecasts then become a linear extrapolation of the current estimate of the trend.
Simple exponential smoothing, SES	$\beta=0$, $\phi=0$	$\beta=0$ and $\phi=0$ imply that $b_t=0\ \forall t\geq 0$. It then follows that $\hat{a}_t=(1-\alpha)\hat{a}_{t-1}+\alpha d_t$. This in turn means $\hat{d}_{t,t+k}=\hat{a}_t$. Here we have made explicit the fact that the SES forecast of all future forecasts (k periods ahead) is simply the forecast of the next periods demand. Ignoring the subscript that gives information on which period we are forecasting yields the common SES formula, $\hat{d}_t=(1-\alpha)\hat{d}_{t-1}+\alpha d_t$.
Naïve forecasting	$\alpha=1$, $\beta=0$, $\phi=0$	This is easy to see from the exponential smoothing formula as this parameter set results in $\hat{d}_{t,t+k}=d_t$.

Table 1. Three popular forecasting methods encapsulated within the DT method

One of our main methodological tools is transfer functions. Transfer functions are useful for studying linear dynamic systems, as they allow convolution in the time domain to be replaced by simple algebra in the complex frequency domain. In the frequency domain there is also a wide range of tools developed by control engineering theorists for understanding the dynamic behaviour of such systems, Nise (2004). The z-transform is often used to solve discrete time difference equations. It is a relatively simple task to develop a block diagram of (1) and manipulate it to obtain the z-transform transfer function of the DT forecasting mechanism (see Figure 1). We refer interested readers to Nise (2004) for information on how to construct and manipulate block diagrams. In Figure 1, the system input is the unit impulse response $\varepsilon(z)=1$ which is the z-tranform of $\varepsilon_t=\{1 \text{ if } t=1,\ 0 \text{ otherwise}$. $D(z)$ is the z-transform of demand series d_t, which obtained by using the z-transform operator $D(z)=\sum_{t=0}^{\infty}(d_t)z^{-t}$. The system output function, $\hat{D}_k(z)$, is the z-transform of the k-period ahead forecast, $\hat{d}_{t,t+k}$.

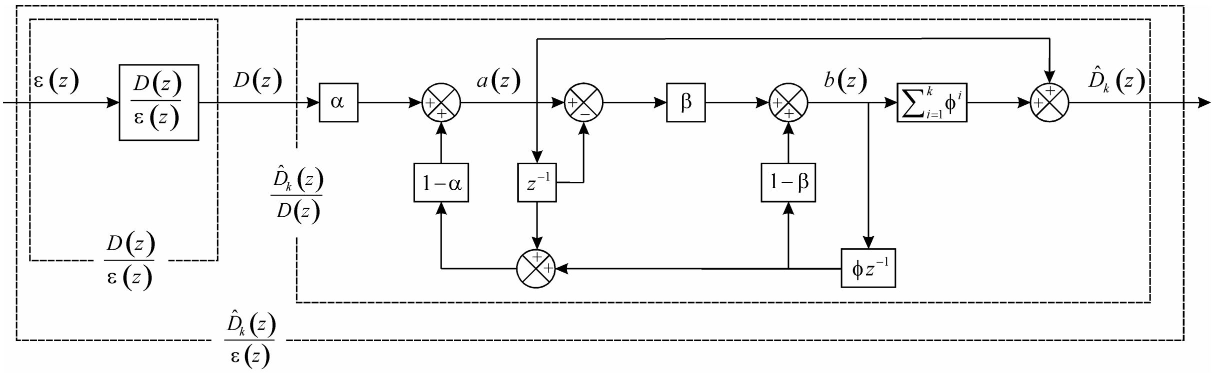

Figure 1. Block diagram of the Damped Trend forecasting method

The discrete time transfer function of (1) is given by the transfer function of the system responding to the unit impulse response. These can be obtained from Figure 1 by assuming that $D(z)/\varepsilon(z)=1$.

$$\frac{\hat{D}_k(z)}{\varepsilon(z)}=\frac{a(z)}{\varepsilon(z)}+\gamma(k,\phi)\frac{b(z)}{\varepsilon(z)}, \tag{2}$$

where $a(z)/\varepsilon(z)$ and $b(z)/\varepsilon(z)$ are the transfer functions of the level $\hat{a}_t$ and the trend component $\hat{b}_t$ accordingly. These are given by

$$\frac{a(z)}{\varepsilon(z)}=\frac{z\alpha\left(z+\phi(\beta-1)\right)}{\phi(1-\alpha)+z(\alpha-\phi-1+\alpha\beta\phi)+z^2}, \tag{3}$$

and

$$\frac{b(z)}{\varepsilon(z)}=\frac{z\alpha\beta(z-1)}{\phi(1-\alpha)+z(\alpha-\phi-1+\alpha\beta\phi)+z^2}. \tag{4}$$

Using (3) and (4), (2) can be written as

$$\frac{\hat{D}_k(z)}{\varepsilon(z)}=\frac{a(z)}{\varepsilon(z)}+\gamma(k,\phi)\frac{b(z)}{\varepsilon(z)}=\frac{z^2\alpha\left(1+\beta\gamma(k,\phi)\right)+z\alpha\left(\phi(\beta-1)-\beta\gamma(k,\phi)\right)}{z^2+z(\alpha-1-\phi+\alpha\beta\phi)+\phi(1-\alpha)} \tag{5}$$

where we have expressed the equation in standard form as coefficients of powers of the *z*-transform operator *z*. From (5), (3) and (4) it is easy to notice that the forecast, the level, and the trend from a second-order system, as the highest power of *z* in the denominator is two. This indicates that complex poles may exist in the system. From a forecasting perspective, complex poles suggest fluctuations in the forecasting errors may exist, even if there are no fluctuations in the demand. From a control theory point of view, complex poles mean oscillations in the system output are created, even by a non-oscillatory input. The subject of the location of the poles is related to a fundamental aspect of dynamic systems, stability - the subject of the next section.

3. Stability of Damped Trend forecasts via Jury's Inners Approach

There are two definitions of stability in the literature. Stability in time series literature is related to stationarity (Box et al., 2008) and is not relevant here. We refer to stability from a control theory perspective (Nise, 2004). Specifically, a stable system will react to a finite input and return to steady state conditions in a finite time. It will return back to steady state in either an over-damped "exponential" response, or with under-damped oscillations that decay away. An unstable system will either diverge exponentially to positive or negative infinity or oscillate with ever increasing amplitude. Oscillations in the forecasts create oscillations in the replenishment order placed in the supply chains and these are costly. So, as a first step to dynamically designing a supply chain replenishment rule, we must first ensure that the forecasting system used within the replenishment rule is stable, Wang et al. (2013).

The stability of the DT method had been investigated in the literature. Although Gardner and McKenzie (1985) promote $0 \le \phi \le 1$, they provide details of the stability region for all $\phi \ge 0$ (but not for the case when $\phi < 0$). Roberts (1982) mentions that a negative ϕ is possible, but the stability criteria were not presented. Hyndman et al. (2008) also emphasise the value of unconventional parameter values. Therefore, in this section, the stability issue will be revisited and the complete stability region will be revealed.

The derivation of the stability criteria for DT forecasting is given in the Appendix. Figure 2 illustrates all possible stability boundaries and how they change for different ϕ. $\alpha\left(1+\phi(\beta-1)\right)>0$ (from Equation (26) in the Appendix) splits the $\{\alpha,\beta\}$ parametric plane into quarters along the lines given by $\alpha=0$ and $\beta=(\phi-1)/\phi$. (27) yields the criteria $2-\alpha-\phi(\alpha-2+\alpha\beta)>0$ which divides the $\{\alpha,\beta\}$ parametric plane along the curve $\beta=(2-\alpha+2\phi-\alpha\phi)/\alpha\phi$, which has an asymptote at $\alpha=0$. From (28), the criteria Δ_{n-1}^{+} divides the parametric plane along $\alpha=(\phi-1)/\phi$, Δ_{n-1}^{-} divides the plane along $\alpha=(\phi-1)/\phi$. The dark grey area in Figure 2 indicates where the stable parameter settings induce conjugate complex poles in the forecasting system. Complex poles within the stability region mean a number of oscillations will be present in the impulse response of the system. Although Figure 2 is a characterisation of the parameter plane, all the boundaries of stability region and the oscillation region (the stability region with complex poles) have been made explicit in Figure 2.

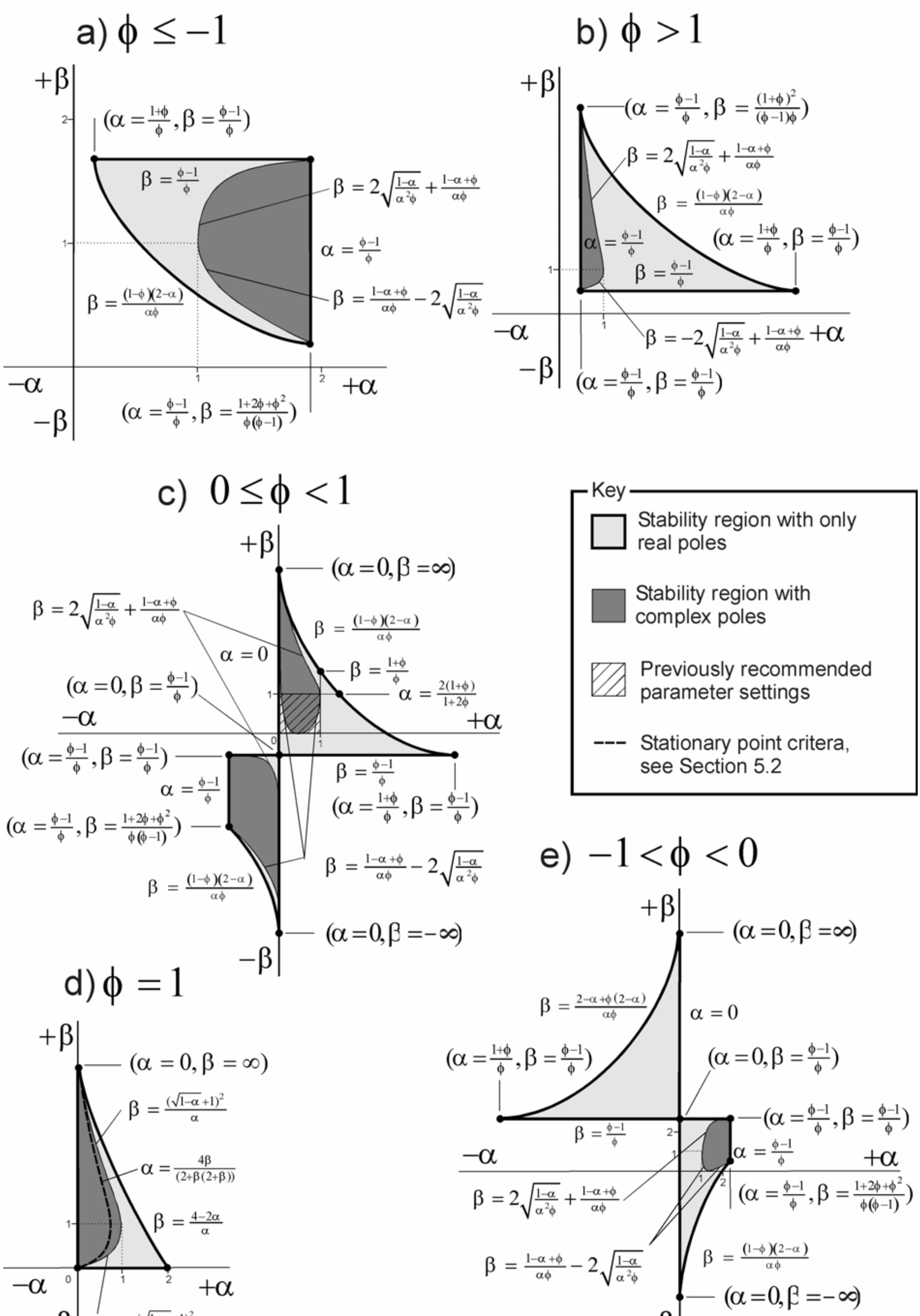

Figure 2. The Damped Trend stability region

When $\phi > 0$, our results are the same as the stability region in Gardner and McKenzie (1985). Similar findings were observed from Figure 2 for Holts method and SES. When $\phi = 1$, we have the Holts method, and the stability conditions are $0 < \alpha < 2$, $0 < \beta < (4 - 2\alpha)/\alpha$. Both the stability region and the oscillation region are identical to McClain and Thomas (1973). When $\beta = 0$ and $\phi = 0$, the SES stability boundary observed $(0 < \alpha \le 2)$, is consistent with the result in Brenner et al. (1968). It is common practice in exponential smoothing models to restrict the smoothing parameters $\{\alpha, \beta\}$ to the $[0,1]$ interval (Holt, 2004; Winters, 1960). The damping parameter has also been proposed to be restricted to $0 \le \phi \le 1$ (Gardner, 1985, 1990; Gardner and McKenzie, 1985, 1989). However, we note that there are stable DT forecasts for a much broader range of parameter values than those usually recommended in the literature. This is interesting as any stable parameter sets produce feasible forecasts and these somewhat unconventional parameter settings may have interesting dynamic properties that may be useful from a production planning and inventory control perspective. In the next section we will integrate the DT forecasting methods with a common replenishment policy so as to investigate this possibility.

4. Integrating Damped Trend forecasting into the Order-Up-To replenishment policy

Having investigated the stability properties of the DT forecasting method we will now look at a single echelon inventory control system where this forecasting technique is used in combination with a linear version of the Order-Up-To replenishment policy (Veinott, 1965; Johnson and Thompson, 1975). We will then derive the transfer functions of the DT / OUT system and use these transfer functions to investigate certain supply chain orientated performance criteria. Specifically we are interested in the bullwhip effect measured by the ratio of the order variance to the demand variance.

4.1. The Order-Up-To policy

A single retailer first receives goods in each period *t*. He observes and satisfies customer demand within the replenishment period, d_t. Any unfilled demand is backlogged. The retailer observes his inventory level and places a replenishment order, o_t, at the end of each period. There is a fixed lead-time of $T_p \in \aleph^0$, between placing an order and receiving that order into stock. We assume that the retailer follows the linear OUT inventory policy. In this policy, orders are placed to raise the inventory position ip_t up to an OUT level or base stock level s_t,

$$o_t = s_t - ip_t. \tag{6}$$

The inventory position is the amount of inventory on-hand + inventory on-order − backlog. The amount of inventory on-hand minus the backlog is known as the net stock ns_t level. The inventory on-order is also known as the Work-In-Progress (WIP), wip_t. The inventory position at time period *t*, ip_t is given by

$$ip_t = ns_t + wip_t, \tag{7}$$

where the net stock is governed by the inventory balance equation, $ns_t = ns_{t-1} + o_{t-T_p-1} - d_t$ and the WIP is governed by $wip_t = wip_{t-1} + o_{t-1} - o_{t-T_p} = \sum_{i=1}^{T_p} o_{t-i}$. The OUT level, s_t , is often estimated from previously observed demand. It can be written as

$$s_t = tns + \hat{d}_{t,t+T_p+1} + \underbrace{\sum_{i=1}^{T_p} \hat{d}_{t,t+i}}_{dwip_t}, \tag{8}$$

where $\hat{d}_{t,t+T_p+1}$ is the forecasted demand in period $t + T_p + 1$ made in period *t*. The Target Net Stock, *tns*, is a safety stock used to ensure a strategic level of inventory availability. *tns* is a time invariant constant. Under the assumptions of piece-wise linear convex inventory holding (*h*) and backlog costs (*b*) then it is common to assume $tns = \sigma_{ns}\Phi^{-1}\left[\frac{b}{b+h}\right]$ (Axsäter, 2006). Here σ_{ns} is the standard deviation of the net stock levels and $\Phi^{-1}[x]$ is the inverse of the cumulative normal distribution function evaluated at *x*. When $tns = \sigma_{ns}\Phi^{-1}\left[\frac{b}{b+h}\right]$, the expected inventory costs per period $I_{£} = \sigma_{ns}(b+h)\varphi[z]$, where $\varphi[\cdot]$ is the probability density function of the normal distribution. The time varying Desired Work In Progress, $dwip_t = \sum_{i=1}^{T_p} \hat{d}_{t,t+i}$ is the sum of the forecasts, made at time *t* of the demand in the periods from $t+1$ to $t+T_p$. The order decision can be rewritten as

$$o_t = tns + \hat{d}_{t,t+T_p+1} + dwip_t - (wip_t + ns_t) = s_t - s_{t-1} + d_t \,. \tag{9}$$

Based upon (8) and the last expression in (9), the *z*-transform transfer function for the order rate, expressed in a manner in which the forecasting system has been left unspecified, is given by

$$\frac{O(z)}{\varepsilon(z)} = \left(1 - z^{-1}\right)\underbrace{\left(\frac{\hat{D}_{T_p+1}(z)}{\varepsilon(z)} + \frac{DWIP(z)}{\varepsilon(z)}\right)}_{\text{z-transform of } s_t} + 1\,. \tag{10}$$

(10) can be seen as a generic transfer function of the DT / OUT system and is a useful departure point for further analysis as the transfer function of the forecasting components can be simply “slotted” into $\hat{D}_{T_p+1}(z)/\varepsilon(z)$ and $DWIP(z)/\varepsilon(z)$ to yield transfer functions of OUT policies with specific forecasting methods. We notice from (8) and (9) that the OUT policy requires two forecasts. One of these forecasts is a prediction, made at time *t* of the demand in the period $t + T_p + 1$. Adapting the DT forecast (see (1)) to achieve this is done by letting $k = T_p + 1$,

$$\hat{d}_{t,t+T_p+1} = \hat{a}_t + \hat{b}_t \sum_{i=1}^{T_p+1} \phi^i = \hat{a}_t + \hat{b}_t \frac{\phi\left(\phi^{T_p+1} - 1\right)}{\phi - 1} = \hat{a}_t + \hat{b}_t \gamma\left(T_p, \phi\right). \tag{11}$$

The other forecast required by the OUT policy is a prediction, made at time t, of the demand over the lead-time. That is, the sum of demand in periods $\left[t+1, t+2, ..., t+T_p\right]$. In the time domain this is

$$dwip_t = \hat{d}_{t,\left[t+1,t+T_p\right]} = \sum_{i=1}^{T_p} \hat{d}_{t,t+i} = \hat{a}_t T_p + \hat{b}_t \eta\left(T_p, \phi\right). \tag{12}$$

where $\eta\left(T_p,\phi\right) = \sum_{j=1}^{T_p}\sum_{i=1}^{j}\phi^i = \left(\frac{\phi}{1-\phi}\right)\left(T_p - \gamma\left(T_p,\phi\right)\right) = \frac{\phi\left(T_p(1-\phi)+\phi\left(\phi^{T_p}-1\right)\right)}{(\phi-1)^2}$. The transfer functions of the DT forecast and WIP target can be built up from the two auxiliary transfer functions, $a(z)/\varepsilon(z)$ and $b(z)/\varepsilon(z)$ previously given in (3) and (4). The z-transforms of the two DT forecasts required by the OUT policy are

$$\begin{aligned}\frac{\hat{D}_{T_p+1}(z)}{\varepsilon(z)} &= a(z) + \gamma\left(T_p+1,\phi\right)b(z) = a(z) + \left(\gamma\left(T_p,\phi\right)+\phi^{k+1}\right)b(z)\\ &= \frac{z^2\alpha\left(1+\beta\left(\gamma\left(T_p,\phi\right)+\phi^{T_p+1}\right)\right) + z\alpha\left(\phi(\beta-1) - \beta\left(\gamma\left(T_p,\phi\right)+\phi^{T_p+1}\right)\right)}{z^2 + z\left(\alpha\left(\beta\phi+1\right)-\phi-1\right)+\phi(1-\alpha)},\end{aligned} \tag{13}$$

and

$$\begin{aligned}\frac{DWIP(z)}{\varepsilon(z)} &= a(z) + \eta\left(T_p,\phi\right)b(z)\\ &= \frac{z^2\alpha\left(T_p + \beta\eta\left(T_p,\phi\right)\right) + z\alpha\left(T_p(\beta-1)\phi - \beta\eta\left(T_p,\phi\right)\right)}{z^2 + z\left(\alpha\left(\beta\phi+1\right)-\phi-1\right)+\phi(1-\alpha)}.\end{aligned} \tag{14}$$

Then using (10) we are able to obtain the transfer function of the OUT replenishment orders,

$$\frac{O(z)}{\varepsilon(z)} = 1 + \frac{(z-1)\alpha\left(\left(1+T_p\right)\left(z+\phi(\beta-1)\right) + \beta(z-1)\left(\gamma\left(1+T_p,\phi\right)+\eta\left(T_p,\phi\right)\right)\right)}{z^2 + z\left(\alpha\left(\beta\phi+1\right)-\phi-1\right)+\phi(1-\alpha)} \tag{15}$$

It can be noticed from (13), (14) and (15) that the denominator of the transfer functions of the DT / OUT system has the same two poles as the transfer function of the forecast system, (5). This implies that the same stability conditions hold for the DT / OUT system as for the Damped Trend system. The nominator of the DT / OUT transfer function is also of second degree, but the coefficients differ. This implies that the zeros of the transfer function lie in a different location to the zeros of the forecasting system.

From (14) the transfer function of the three special DT variants Holts Method, SES and Naïve forecasting can be easily derived and are given in Table 2. As the trend component of the demand process is not explicitly forecasted in the SES and the Naïve forecasting models, we follow the commonly adopted procedure and set the DWIP term to the product of the lead-time and the most recent forecast (see for example Chen et al., 2000; Lalwani et al., 2006).

Forecasting Method	Transfer Functions of $\hat{D}_{T_p+1}$ and *DWIP*
Holts method	$\frac{\hat{D}_{T_p+1}(z)}{\varepsilon(z)}=\frac{z^2\alpha\left(1+T_p\beta+\beta\right)-z\alpha\left(1+T_p\beta\right)}{z^2+z\left(\alpha\left(1+\beta\right)-2\right)+1-\alpha}$ $\frac{DWIP(z)}{\varepsilon(z)}=\frac{z^2\alpha T_p\left(2+T_p\beta+\beta\right)+z\alpha T_p\left(\beta\left(1-T_p\right)-2\right)}{2\left(z^2+z\left(\alpha\left(1+\beta\right)-2\right)+1-\alpha\right)}$
SES	$\frac{\hat{D}_{T_p+1}(z)}{\varepsilon(z)}=\frac{z\alpha}{z+\alpha-1}$; $\frac{DWIP(z)}{\varepsilon(z)}=\frac{z\alpha T_p}{z+\alpha-1}$
Naïve forecasting	$\frac{\hat{D}_{T_p+1}(z)}{\varepsilon(z)}=1$; $\frac{DWIP(z)}{\varepsilon(z)}=T_p$

Table 2. Transfer functions of $\hat{D}_{T_p+1}$ and *DWIP* for the special cases of DT / OUT system

5. Bullwhip behaviour of DT / OUT system and its special cases

5.1. The Bullwhip effect in the frequency domain

A well-known criterion to indicate a bullwhip effect is the ratio of the order variance to demand variance. The order transfer function is important as it contains information on the well-known bullwhip effect, Dejonckheere et al (2003). $O(z)/\varepsilon(z)$ be the order transfer function of the DT / OUT system, reacting to the unit impulse response, ε_t and was given in (15).

Let $\frac{D(z)}{\varepsilon(z)}=d_0+d_1z^{-1}+d_2z^{-2}+d_3z^{-3}+...$ be the transfer function of the demand process. As each d_t is arbitrary, $D(z)/\varepsilon(z)$ could represent any demand pattern for which there is a *z*-transform. This demand pattern could be, for example, stationary, non-stationary, random, deterministic and operating over a finite or an infinite time horizon.

From Tsypkin (1964) the variance of the demand is $\sigma_d^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{D(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 d\omega$ and the variance of the orders when the DT / OUT policy is reacting to that demand is $\sigma_o^2=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{O(e^{j\omega})}{\varepsilon(e^{j\omega})}\frac{D(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 d\omega$. Here $z=e^{j\omega}$, the complex frequency, ω. For bullwhip to exist then $\sigma_o^2-\sigma_d^2>0$ which implies that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{O(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2\left|\frac{D(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 d\omega>\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{D(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 d\omega. \qquad (16)$$

If $\left|\frac{O(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 > 1\ \forall\omega$ then a bullwhip effect always exists regardless of $\left|\frac{D(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|\ \forall\omega$. This is a sufficient condition for bullwhip. Notice that in order to cope with non-stationary demand processes we have used differences between the two variances as an indicator of whether a bullwhip effect exists. This measure was introduced by Gaalman and Disney (2012). When $\sigma_o^2 - \sigma_d^2 > 0$ a bullwhip effect exists, when $\sigma_o^2 - \sigma_d^2 < 0$ bullwhip is avoided. This is somewhat different to the usual variance ratio $\left(\sigma_o^2/\sigma_d^2\right)$ that is used to indicate a bullwhip effect exists when $\sigma_o^2/\sigma_d^2 > 1$. The variance ratio has not been used as it is unable to distinguish bullwhip behaviour when non-stationary demands are present.

In order to simplify notation we will let $\left|\frac{O(e^{j\omega})}{\varepsilon(e^{j\omega})}\right|^2 = AR$, for Amplitude Ratio as it represents the amplification of the individual harmonics with a frequency of ω radians per period, Nise (2004). If the $AR < 1$ holds only for a certain range of harmonic frequencies, it is still possible (but not certain) that $\sigma_o^2 < \sigma_d^2$ - it depends on the frequency characteristics of the demand pattern. Investigating the Amplitude Ratio of different frequencies from frequency response plots, we are able to gain insight into how the forecasting and replenishment system creates or avoids bullwhip for arbitrary demand patterns.

The frequency response (the Amplitude Ratio over ω) of a discrete time system is a function with a periodicity of 2π. However, we only need to study the AR for frequencies in the period $[0,\pi]$, as the frequency response plot on $[-\pi,0]$ is a simple reflection of $[0,\pi]$ about the ordinate. Furthermore we note that $AR|_{\omega=0} = 1$ and $\frac{\mathrm{d}AR}{\mathrm{d}\omega}\big|_{\omega=0} = 0$ for our DT / OUT system. $AR|_{\omega=0} = 1$ because the transform function $O(z)/\varepsilon(z)$ has equal power of z in both the nominator and denominator and it is a passive system with a unity gain. The $\frac{\mathrm{d}AR}{\mathrm{d}\omega}\big|_{\omega=0} = 0$ is required because of the reflection property of the frequency response. By studying the monotonicity, the final value of an AR curve and the number of stationary points in the frequency response, we are able to characterise the frequency response of a system and draw conclusions on bullwhip behaviour. Let us now take a look at the frequency response for the three special cases of the DT / OUT system. First we analyse the three special variants of the DT model (Naïve, SES and Holts) in 5.2 and then move on the to DT situation in section 5.3.

5.2. Frequency response of three special cases of DT / OUT policy (*with Naïve, SES, or Holts forecasts*)

When Naïve forecasts ($\phi = \beta = 0,\ \alpha = 1$) are used in the OUT policy with unspecified lead-time, the *AR* is

$$AR = \sqrt{1 + 2\left(1+T_p\right)\left(2+T_p\right)\left(1-\cos(\omega)\right)}\,, \tag{17}$$

which is strictly increasing in ω within the interval $(0,\pi)$, as

$$\frac{dAR}{d\omega}=\frac{\left(1+T_p\right)\left(2+T_p\right)\sin(\omega)}{\sqrt{1+2\left(1+T_p\right)\left(2+T_p\right)\left(1-\cos(\omega)\right)}}>0\,, \tag{18}$$

$AR|_{\omega=0}=1$ and $AR|_{\omega=\pi}=3+2T_p$, see Figure 3a. This means $AR\geq 1$ for all frequencies. This implies that the OUT policy with Naïve forecasts will produce bullwhip for every possible demand pattern and for all lead-times.

Figure 3b shows when stable SES forecasting $\left(\phi=\beta=0,\ 0<\alpha<2\right)$ is used in the OUT policy. The *AR* is given by

$$AR=\sqrt{\frac{2+\alpha\left(2+\alpha+2T_p\left(2+\alpha+T_p\alpha\right)\right)-2\left(1+T_p\alpha\right)\left(1+\alpha\left(1+T_p\right)\right)\cos\left(\omega\right)}{2+\alpha\left(\alpha-2\right)+2\left(\alpha-1\right)\cos\left(\omega\right)}}\,, \tag{19}$$

which is also a strictly increasing function within the frequency interval $\left(0,\pi\right)$ as the first derivative is

$$\frac{dAR}{d\omega}=\frac{\left(1+T_p\right)\alpha^3\left(2+T_p\alpha\right)\sin\left(\omega\right)}{\sqrt{\left(2+\alpha\left(\alpha-2\right)+2\left(\alpha-1\right)\cos\left(\omega\right)\right)^3\begin{pmatrix}2+\alpha\left(2+\alpha+2T_p\left(2+\alpha\left(1+T_p\right)\right)\right)-\\ 2\left(1+T_p\alpha\right)\left(1+\alpha\left(1+T_p\right)\right)\cos(\omega)\end{pmatrix}}}>0\,, \tag{20}$$

when $T_p\geq 0$, $0\leq\alpha\leq 2$ and $0<\omega<\pi$. Together with $AR|_{\omega=0}=1$ we can deduce that $AR\geq 1$ $\forall\omega$. In another words, the OUT policy with SES forecasting will always produce bullwhip effect for all demand patterns and for all lead-times. This finding is consistent with the results in Dejonckheere et al. (2003), but their conjecture was not formally proved.

When stable Holts forecasting mechanism $\left(\phi=1,\ 0<\alpha<2,\ 0<\beta<(4-2\alpha)/\alpha\right)$ is used, the AR is given by

$$AR=\sqrt{\frac{\begin{pmatrix}2+\alpha\left(1+2T_p\right)+\alpha\beta T_p\left(2+T_p\right)-\\ \left(2+\alpha\left(1+2T_p\right)+\alpha\beta\left(1+T_p\right)^2\right)\cos\left(\omega\right)\end{pmatrix}^2+\alpha^2\left(1+\beta\left(1+T_p\right)\right)^2\sin\left(\omega\right)^2}{\left(\alpha\left(1+\beta\right)-2-\left(\alpha-2\right)\cos\left(\omega\right)\right)^2+\alpha^2\sin\left(\omega\right)^2}}\,. \tag{21}$$

The frequency response originates at $AR|_{\omega=0}=1$ and ends at

$$AR|_{\omega=\pi}=\frac{4+\alpha\left(2+\beta+2T_p\left(2+\beta\left(2+T_p\right)\right)\right)}{4-\alpha\left(2+\beta\right)}>1, \tag{22}$$

see Figure 3c. By investigating the derivative

$$\frac{dAR}{d\omega}=\frac{\begin{pmatrix}4\left(1+T_p\right)\alpha^2\left(2+\beta\left(1+T_p\right)\right)\left(4+T_p\alpha\left(T_p\beta+2\left(1+\beta\right)\right)\right)\\ \left(\alpha\left(1+\beta\left(1+\beta\right)\right)-2\beta-\left(\alpha\left(1+\beta\right)-2\beta\right)\cos\left(\omega\right)\right)\cos\left(\tfrac{\omega}{2}\right)\sin\left(\tfrac{\omega}{2}\right)^3\end{pmatrix}}{\sqrt{\begin{matrix}\left(\left(\alpha\left(1+\beta\right)-2-\left(\alpha-2\right)\cos\left(\omega\right)\right)^2+\alpha^2\sin\left(\omega\right)^2\right)^3\\ \left(\begin{pmatrix}2+\alpha\left(\left(1+2T_p\right)+T_p\beta\left(2+T_p\right)\right)-\\ \left(2+\alpha\left(\left(1+2T_p\right)+\beta\left(1+T_p\right)^2\right)\right)\cos\left(\omega\right)\end{pmatrix}^2+\alpha^2\left(1+\beta\left(1+T_p\right)\right)^2\sin\left(\omega\right)^2\right)\end{matrix}}}, \tag{23}$$

we find that in between these two points there are two different *AR* responses. Either the *AR* is strictly increasing in ω or, when $\alpha<4\beta/\left(2+\beta\left(2+\beta\right)\right)$, there is a stationary point within the $\omega\in\left(0,\pi\right)$ interval, see Figure 2d. The stationary point criteria, $\alpha<4\beta/\left(2+\beta\left(2+\beta\right)\right)$, was found by setting $\mathrm{d}AR/\mathrm{d}\omega=0$ and solving for α. Interesting there is no influence of T_p on this boundary. When there is a stationary point present in $\omega\in\left(0,\pi\right)$, the *AR* is an increasing function in ω until $\omega=\arccos\left(\left(\alpha+\beta\left(\alpha\left(1+\beta\right)-2\right)\right)/\left(\alpha\left(1+\beta\right)-2\beta\right)\right)$ at which point it becomes a decreasing function until $\omega=\pi$. The stationary point, if it exists, will be a maximum. Therefore, as $AR\geq 1\ \forall\omega$, and bullwhip is generated by this system for any lead-time, any demand pattern.

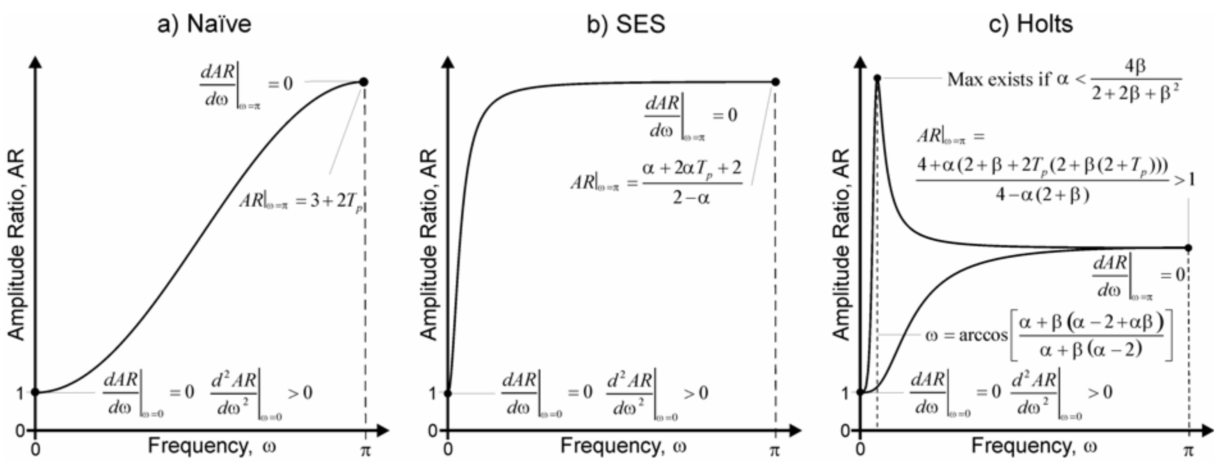

Figure 3. Frequency response of the OUT policy with Naïve, SES and Holts forecasts

Using these facts we have proved that three special cases of DT / OUT system – the Order-Up-To with Naïve, SES and Holts forecasts, will always generate bullwhip for any demand pattern and all lead-times.

5.3. Frequency response of the OUT Policy with Damped Trend forecasts.

The DT / OUT frequency response is more complex than the previous cases. The conditions for which the *AR* curve is greater than one cannot be determined analytically. Because of this we now divide the investigation into three steps / parts. First, we analyse the low-frequency response (ω near 0). Second, we look at the high frequency response (ω at π). Third, we determine the number of stationary points within the interval $(0,\pi)$ in order to gain information on the minima and maxima and by this obtain insights on the fluctuating behaviour of the frequency response. In this section of the paper, many of the required equations (that we have obtained via Mathematica ® Wolfram Research) have been omitted as they are lengthy, complex and difficult to analyse.

We know that $AR|_{\omega=0}=1$ and $\frac{\mathrm{d}AR}{\mathrm{d}\omega}\big|_{\omega=0}=0$. The sign of the second derivative of the *AR* function at $\omega=0$ indicates a local increasing or decreasing behaviour in the AR function near zero. This is an indicator of whether a bullwhip effect is produced when demand is a dominated by low frequency harmonics. Because of this, the geometrical implications of $\frac{\mathrm{d}^2AR}{\mathrm{d}\omega^2}\big|_{\omega=0}$ determines whether the *AR* value when ω is near 0 (denoted as $AR|_{w\to 0}$) can be greater or smaller than 1. We also investigate high-frequency responses (ω near π) from the value of *AR* at $\omega=\pi$. Finally, the shape of the *AR* curve is determined by combining information on the number of stationary points within the $\omega\in(0,\pi)$ interval with information obtained from the first two steps. We can then characterise the *AR* plot for the DT / OUT system and gain insights on bullwhip behaviour.

First, consider low-frequency behaviour. Although $AR|_{\omega=0}=1$ and $\frac{\mathrm{d}AR}{\mathrm{d}\omega}\big|_{\omega=0}=0$, the second derivative can be positive, zero or negative. The sign of the second derivative has geometrical implications. If $\frac{\mathrm{d}^2AR}{\mathrm{d}\omega^2}\big|_{\omega=0}>0$ the graph of *AR* will be *convex* in a small interval $[0,\delta)$ with a local minimum at $\omega=0$, which means $AR|_{0<\omega<\delta}>1$. If $\frac{\mathrm{d}^2AR}{\mathrm{d}\omega^2}\big|_{\omega=0}<0$, the *AR* curve will be *concave* in a small interval $[0,\delta)$ and the point at $\omega=0$ is a local maximum, which means $AR|_{0<\omega<\delta}<1$. A *convex AR* indicates that the pure DT / OUT system will generate bullwhip when demands are dominated by the low frequency harmonics in $[0,\delta)$. A *concave AR* implies that the DT / OUT policy is able to avoid generating bullwhip for demands that are dominated by low frequency harmonics in $[0,\delta)$. If the second derivative is zero, theoretically, the origin could be an inflection point if the lowest-order non-zero derivative is of an odd order since the first derivative is already zero. However, for all of the possible DT settings, we found the lowest-order non-zero derivative is always of even order. Then, concurring with fundamental knowledge of the periodicity of the frequency response, the stationary point at $\omega=0$ has to be either a local maximum or a local minimum, when the second derivative is zero.

When the lead-time $T_p=1$ and $\phi\in[-1,0)\cup(0,1)$ (that is, if $-1\le\phi<0$ or $0<\phi<1$, see Figure 2c) or 2e)) then $\frac{\mathrm{d}^2AR}{\mathrm{d}\omega^2}\big|_{\omega=0}>0$. This means that these settings will always amplify low frequency harmonics. When $\phi>1$ or $\phi\le-5$ (plots 4a and 4d) the *AR* near $\omega=0$ is always *concave*, implying low frequency harmonics are attenuated. However when $\phi<-1$, the second derivative can be positive, negative or zero. Figure 4 maps out the areas of the parametric plane where bullwhip is avoided for demand patterns that are dominated by low frequency harmonics. The curves which separate out the two classes of bullwhip behaviour for when $-3<\phi<-1$ are $\alpha=\frac{\phi-1}{\phi}$ and $\beta=\left\{\frac{-2-\alpha}{\alpha\phi},\frac{\phi-1}{\phi}\right\}$. When $-5<\phi\le-3$ the curves $\alpha=\frac{\phi-1}{\phi}$

and $\beta = \frac{(1+\phi)(2-\alpha)}{\alpha\phi}$ separate the different classes of bullwhip behaviour. These curves were all obtained by setting $\left.\frac{\mathrm{d}^2 AR}{\mathrm{d}\omega^2}\right|_{\omega=0} = 0$ and solving for the relevant variables.

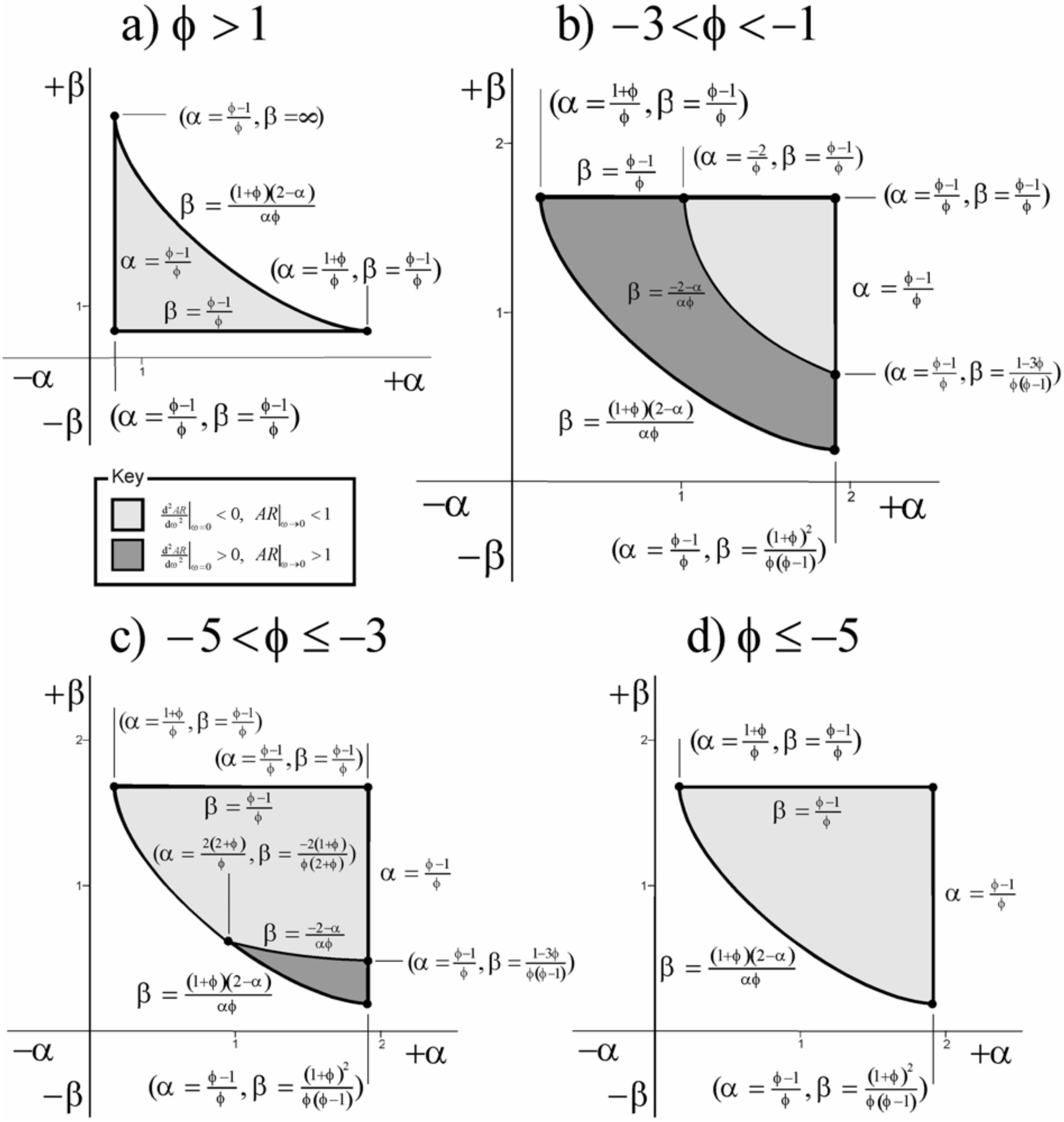

Figure 4. Possible settings that result in $AR<1$ near $\omega=0$ when $T_p=1$

Second, we consider the high-frequency bullwhip behaviour near $\omega=\pi$ when $T_p=1$. DT forecasts with $\phi>1$ or $\phi\le-3$ always generate bullwhip for high-frequency demands as $AR|_{\omega=\pi}>1$. If $\phi=-1$, then $AR|_{\omega=\pi}=1$. Interesting, here are some circumstances when $AR|_{\omega=\pi}<1$, see Figure 5. $AR|_{\omega=\pi}<1$ indicates that the DT enabled OUT policy is able to avoid inducing bullwhip when demand is dominated by high frequency harmonics. This bullwhip avoidance occurs for: $0<\phi<1$ when $\frac{\phi-1}{\phi}<\alpha<0$ and $\frac{\phi-1}{\phi}<\beta<\frac{-1}{\phi}$; $-1<\phi<0$ when $0<\alpha<\frac{\phi-1}{\phi}$ and $\frac{-1}{\phi}<\beta<\frac{\phi-1}{\phi}$; and for $-3<\phi<-1$ when $\alpha<\frac{\phi-1}{\phi}$ and $\frac{-(2+\alpha)(1+\phi)}{\alpha\phi(1+2\phi)}<\beta<\frac{-1}{\phi}$.

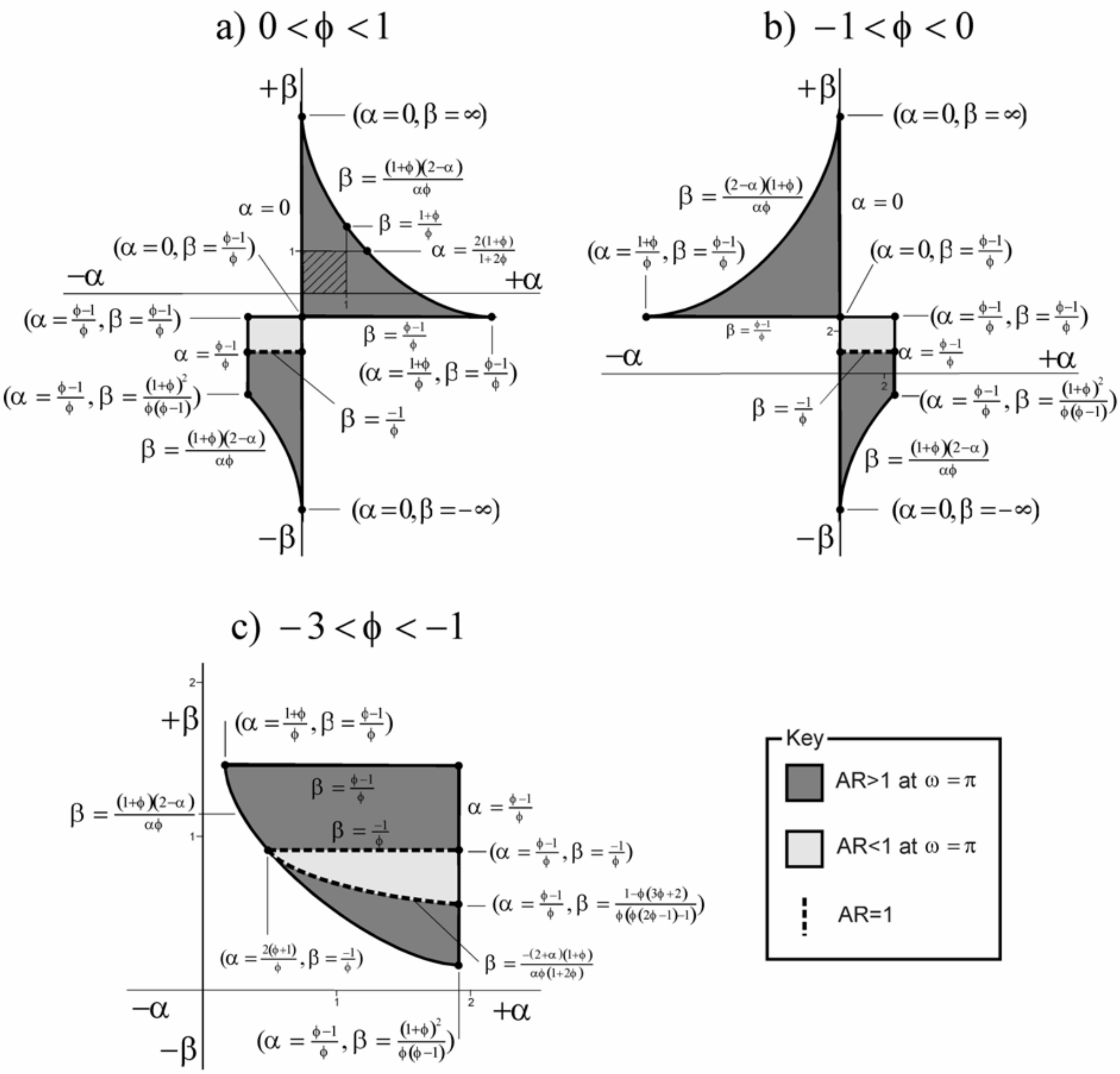

Figure 5. Possible settings that result in $AR\le 1$ near $\omega=\pi$ when $T_p=1$

Third, we investigate the number of stationary points within $\omega\in(0,\pi)$ to determine the shape of the *AR* curve. Figure 6 illustrates how many possible stationary points there are for different stable parameter settings when $T_p=1$. The numbers $\{0,1,2\}$ in Figure 6 denote the possible number of stationary points in each region. The fact that there are only a maximum of two stationary points in the frequency domain should come as no surprise as we are dealing with a second order system (a system with two poles and two zeros). The number of stationary points within each region was found by carefully counting the number of solutions to $\frac{dAR}{d\omega}=0$. We regret that the equations involved in this task are too lengthy and complex to include in this paper. However, with this information, as well as the information about the frequency responses at low and high-frequencies, we can characterise the shape of the *AR* curve. It is easy to verify from our discussion of figures 4 and 5 that $AR|_{\omega\to 0} < 1$ and $AR|_{\omega=\pi} < 1$ cannot exist together (see plot 6e for an alternative verification). This concurs with the fact that the system has two poles and two zeros (i.e. the amplitude has at most two stationary points, a minimum and maximum) which prohibits the existence of a system with $AR|_{\omega\to 0} < 1$ and $AR|_{\omega=\pi} < 1$. It implies that the OUT policy with DT forecasts cannot avoid amplifying individual harmonics for the whole range of the frequency spectrum, though there may or may not be a bullwhip effect. When $0 \le \phi \le 1$, $\alpha > 0$ and $\beta > 0$ (this is the traditionally advised parameter setting by Gardner and McKenzie (1985)) then $AR|_{\omega\to 0} > 1$, $AR|_{\omega=\pi} > 1$, and there is either one stationary point or no stationary points within $\omega \in (0, \pi)$. It is then not difficult to visualise the *AR* curve and realise that *AR* > 1 for all frequencies ω (using the same line of reasoning that we used for the Holts method). This means the traditional DT settings will always result in an OUT policy that produces a bullwhip effect for any demand pattern. Given the relevance in the literature to these parameter settings, this is a important new insight. The same bullwhip behaviour happens at the parameter settings within the dark grey areas of Figure 6. Nevertheless, some of these settings may still provide relatively good performance, see for example b) and e) in Figure 7.

For non-traditional parameter settings however, it is interesting that for some demand patterns the bullwhip effect can be avoided. Some evidence is given in Figure 7 where we have characterised all possible behaviours of the frequency response. The parameter settings of frequency response plot a) and h) in Figure 7 are from the bullwhip avoidance areas in Figure 6a and 6f. Bullwhip is avoided when demand is dominated by low-frequency harmonics, and there are two stationary points when $0 < \omega < \pi$ which is consistent to the information in Figure 6a and 6f. But for the rest of the frequency spectrum, the performance of these two parameter settings (7a and 7f) is not attractive. In plot Figure 7g, the DT / OUT policy performs relatively well both at low-frequencies and high-frequencies, but harmonic amplitudes are magnified for the rest of the frequency interval.

It is impressive that $AR > 1$ for only a very few frequencies in plot c) and d) at Figure 7. For example, in plot c), $AR > 1$ only occurs between $0 < \omega < 0.1$ radians per sample interval, and for the remaining of the frequency spectrum, the amplitude ratio is actually much smaller than unity. These two settings are from the bullwhip avoidance area in Figure 6b and 6c (the grey areas). A similar *AR* plot was noticed by Dejonckheere et al. (2003). However they used exponential smoothing and modified the OUT policy by adding a proportional feedback controller into the inventory and WIP feedback loops to create a discrete time version of the IOBPCS model (Towill, 1982). Our results suggest that the Damped Trend forecasting mechanism could allow the OUT policy to avoid the bullwhip effect without the difficult and expensive modification to the OUT policy advocated by Dejonckheere et al. (2003). This is another important managerial insight.

By repeating our analysis (details omitted for brevity) for different lead-times we find that when $T_p > 1$, for low-frequency demands, the parametrical plane $\phi > 1$ still can enable the DT / OUT system to avoid the bullwhip effect as $AR|_{\omega\to 0} < 1$. For high-frequency demands, when $0 < \phi < 1$ the area of the parametrical plane that is able to avoid the bullwhip effect (in Figure 5a) will become smaller if we increase the lead-time; when $-1 < \phi < 0$, the region within the parametrical plane where bullwhip is avoided (in Figure 5b) changes in a complex manner. It has different shapes when the lead-time changes from an odd number to an even number.

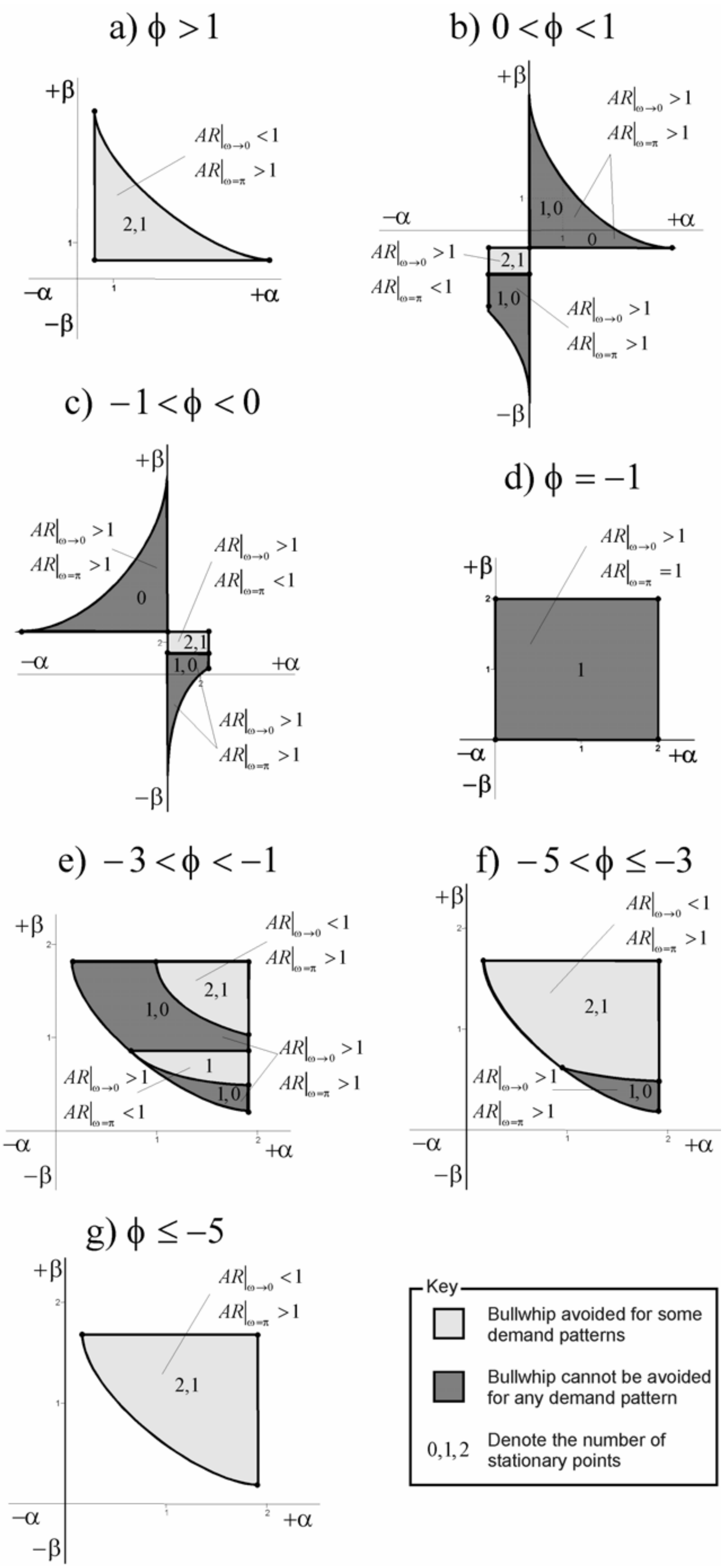

Figure 6. Number of stationary points within $\omega \in (0, \pi)$ and value of *AR* near $\omega = 0$ and $\omega = \pi$ when $T_p = 1$

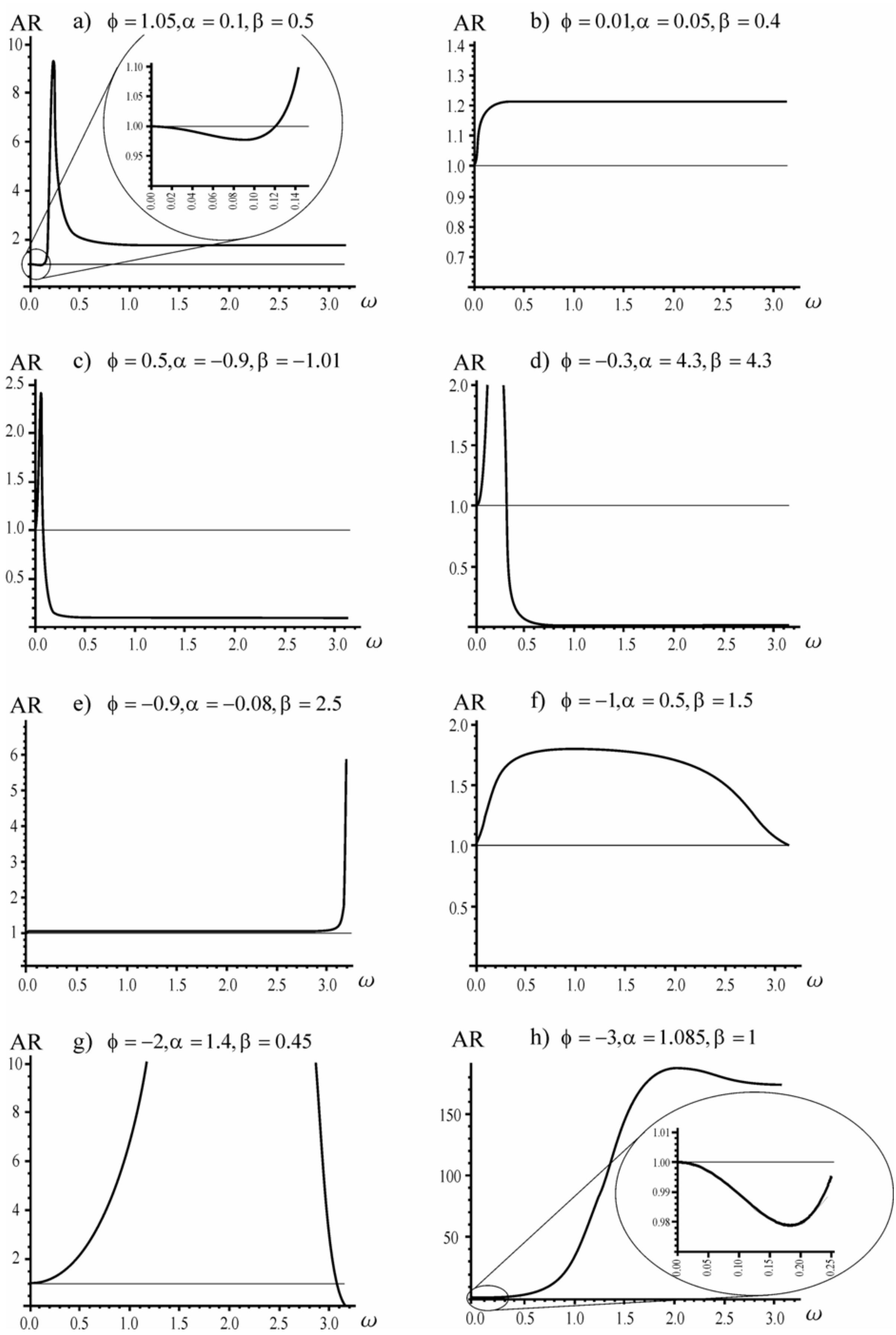

Figure 7. Some examples of the *AR* plot when $T_p = 1$

When $\phi < -1$, bullwhip avoiding areas of the parametrical plane for both low-frequency (Figure 4b, 4c, 4d) and high frequency (Figure 5c) demand will disappear and reappear in sophisticated manners when the lead-time switches between an odd number and an even number.

Although a set of stable parameters that can avoid generating bullwhip over the whole range of frequencies does not exist in the DT / OUT system and the traditional parameter setting always produce bullwhip, the results we have obtained indicate that for some demand patterns the OUT policy with DT forecasting may be able to avoid the bullwhip effect (as measured by the variance ratio) with somewhat unconventional parameter values. This is a type of dynamic behaviour that is not present when Naïve, SES or Holts Method is used as a forecasting method within the OUT policy. Note however that the bullwhip avoidance areas that we have found in the DT / OUT parameter plane are only necessary conditions, so in the next section we will conduct some numerical investigations to verify our findings and gain an understanding of how relevant our necessary conditions are.

6. Numerical verification of the bullwhip avoidance properties in the DT / OUT system

In this section we will verify our analytical insights via simulation. First we will verify the high and low frequency behaviour of the DT / OUT system. Then we will simulate the response to a real life demand pattern to demonstrate the DT / OUT does indeed exhibit bullwhip avoidance behaviour in a practical setting. Then we explore the bullwhip performance of the necessary conditions that we found in section 5 (the area in Figure 6b given by $\left\{0 < \phi < 1,\ \frac{\phi-1}{\phi} < \alpha < 0,\ \frac{\phi-1}{\phi} < \beta < \frac{-1}{\phi}\right\}$) with a grid search. This area was chosen as we previously noted it had a very desirable frequency response, acting like a low-pass filter.

6.1. Verification of the low and high frequency harmonic behaviour of the DT / OUT policy

In this section we will verify that the bullwhip avoidance is possible when demand consists of a single harmonic frequency. This harmonic can be of a low frequency (to verify the bullwhip avoidance areas identified in Figure 4) or a high frequency (to verify the bullwhip avoidance areas identified in Figure 5). To do this we constructed an Excel based simulation of the DT / OUT policy with unit lead-times ($T_p = 1$). Demand was assumed to be made up of a single sine wave with a mean of 10, unit amplitude and a frequency of $\omega \in \{0.02,\ 3.1\}$ radians per period. We determined the bullwhip ratio, σ_o^2 / σ_d^2 from 4000 periods after an initialisation period of 1000 periods in order to avoid any transient responses produced by initial conditions. Sample numerical results are given below in Table 3. These values of α, β and ϕ were chosen because they lie in each of the bullwhip avoidance regions in Figure 4 and 5. They verify our theoretical investigations that the OUT policy with DT forecasting can indeed eliminate the bullwhip effect when demand is a single harmonic frequency.

6.2. Application of our theory to a real-life data set

To verify our theoretical and numerical results in a more realistic / practical situation, we also selected a real-life data set of (consisting of 128 daily demand values) placed on a manufacturer of fast moving consumer goods in the household goods category from a UK

Parameters	Demand Frequency ω (radians per period)						
	0.02				3.1		
α	0.14	1.6	1.1	1.1	-0.5	2	1.4
β	0.14	1.6	1.1	1.1	-1	2	0.45
ϕ	1.1	-1.5	-4.5	-5.5	0.6	-0.6	-2
Bullwhip σ_o^2/σ_d^2	0.9768	0.9964	0.9824	0.9624	0.4278	0.5389	0.1697
Bullwhip avoidance area in Figure	4a	4b	4c	4d	5a	5b	5c

Table 3. Numerical results from a 4000 period simulation verifying our theoretical results

supermarket chain This is the same demand pattern used in Dejonckheere et al (2003). Interestingly it contains an upward trend followed by a downward trend and another upward trend. In the middle of this sequence there is also a significant short-lived spike in demand. We assumed the lead-time $T_p = 1$, which matched the real-life scenario.

The original demand pattern is decomposed into 64 harmonic frequencies via the Fast Fourier Transform (FFT). The FFT is a computational tool which facilitates signal analysis and efficiently computes the Discrete Fourier Transform of a time series. It can be found as a built-in function in Microsoft Excel. It is an easy task to calculate the amplitude and phase of each harmonic frequency and identify dominant frequencies in the spectrum. Figure 8 shows the amplitude of the harmonic cosine waves that make up the original demand pattern. We see that there is dominant frequency of $\frac{\pi}{64}$ radians per period but there also a number of significant frequencies across the whole range. This information suggests that we are looking for an OUT policy frequency response that has a similar shape to plot c) in Figure 7. This occurs in the region of $\left\{0<\phi<1,\ \frac{\phi-1}{\phi}<\alpha<0,\ \frac{\phi-1}{\phi}<\beta<\frac{-1}{\phi}\right\}$, the area we previously identified as having the characteristics of a low-pass filter. After some exploration of suitable parameter sets in this are we settled $\{\phi=0.077,\ \alpha=-5.695,\ \beta=-12.13\}$ as it had both good bullwhip avoidance behaviour and maintained good control over the inventory levels. From the Excel simulation the bullwhip ratio obtained was $\frac{\sigma_o^2}{\sigma_d^2}=0.767727$. The sum of the product of the amplitudes of the Fourier Transform harmonics (also shown in Figure 8) and the *AR* predicted a bullwhip ratio of 0.766513, thus validating our theoretical approach. Figure 9 gives a time series plot of the demand, the two-period ahead forecast, the production orders and the inventory levels maintained by the DT / OUT system.

6.3 Investigation of the necessary conditions for bullwhip avoidance performance for a real-life demand pattern

Finally, we will now conduct a grid search of the DT / OUT policy with $T_p = 1$ reacting to the real-life demand pattern in Figure 8. The grid search explored the low pass filter area given by $\left\{0<\phi<1,\ \frac{\phi-1}{\phi}<\alpha<0,\ \frac{\phi-1}{\phi}<\beta<\frac{-1}{\phi}\right\}$ for different values of $0.1\le\phi\le 0.9$ in steps of 0.1. The stable range of α and β was divided up into 50 increments. We can see that the necessary conditions, for this particular demand pattern, resulted in bullwhip avoidance for about 20% of the possible cases. Interestingly, in order to avoid bullwhip, we find that α should be selected in the mid-range of its stability region and β should be selected in the high (least

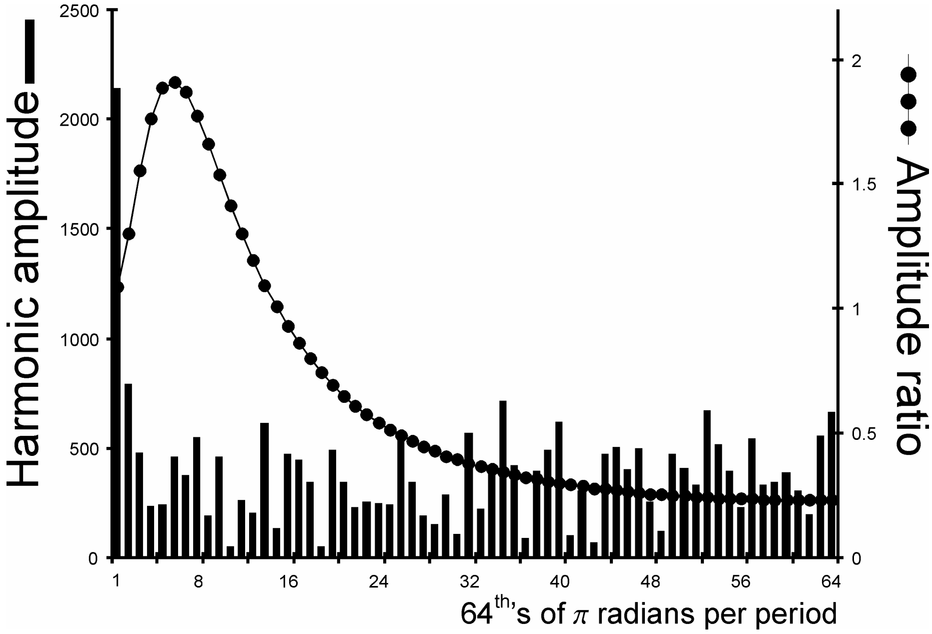

Figure 8. FFT analysis of the demand for a FMCG in the household goods category

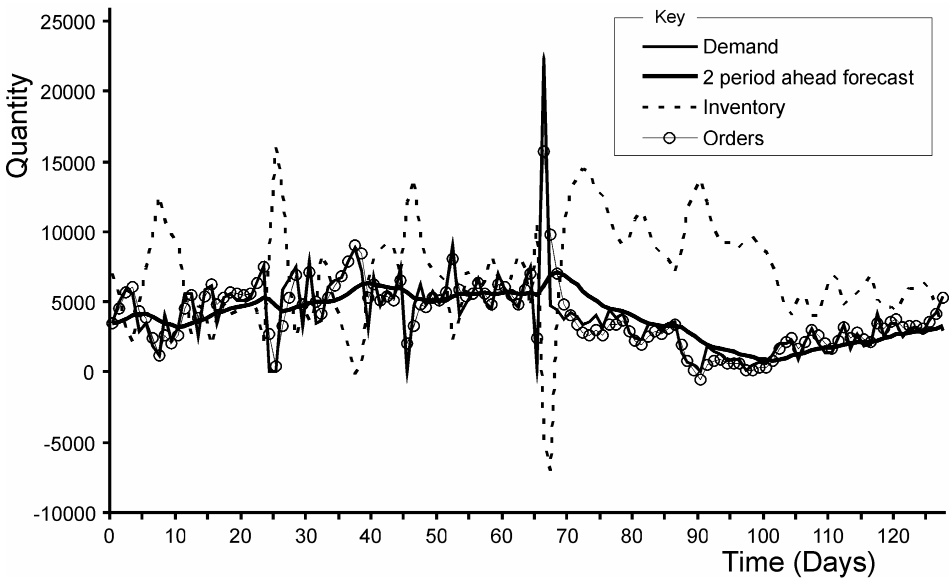

Figure 9. Simulated response by the OUT policy with DT forecasts

negative) end of its stability region. This suggests that the forecasting parameters need to be carefully matched to the demand process.

7. Concluding remarks

Stability issues of DT forecasting mechanism have been previously discussed in the literature. However, we have demonstrated that the DT forecasts are stable over a much broader range of parameter values than is currently acknowledged. We then proceeded to analyse the consequences of using the DT forecasting method within the OUT policy for the complete range of stable forecasting parameters. We have shown the OUT policy with three special cases of DT forecasts - Naïve, SES and the Holts method forecasts, will always generate

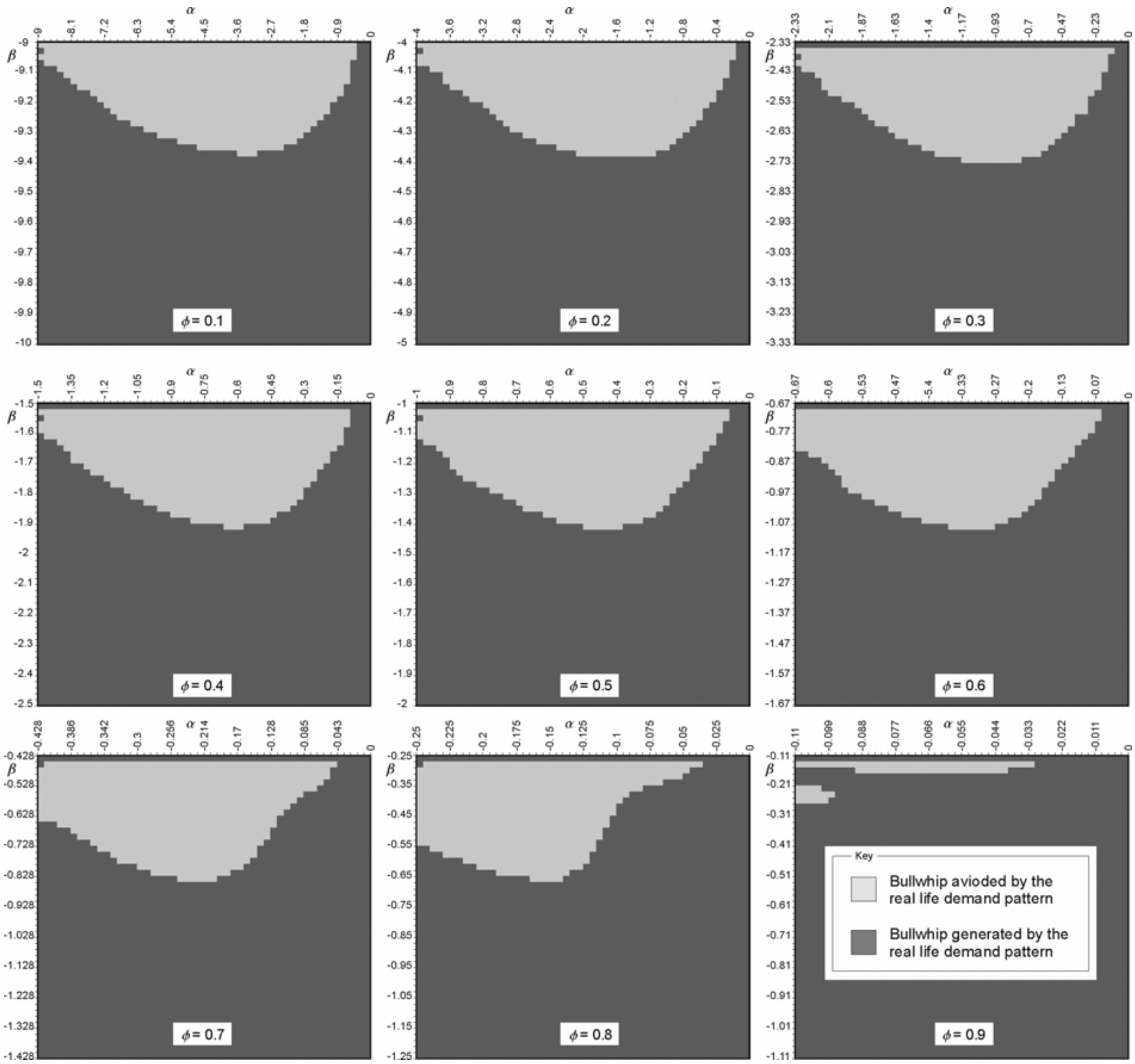

Figure 10. Investigation of the performance of the necessary bullwhip criterion

bullwhip for any demand patterns and for all lead-times. Our findings strengthen, sharpen and refine the arguments of Dejonckheere et al. (2003). However, we have further shown that for some demand patterns the OUT replenishment policy with DT forecasting mechanism is able to avoid generating bullwhip. This is a qualitatively different bullwhip behaviour that is not present with other, more traditional forecasting policies.

We have verified our theoretical insights via a numerical simulation and shown how to apply them to a real life demand pattern. In one real-life demand pattern we were also able to reduce bullwhip generated by the OUT policy to 0.77. This level of performance cannot be achieved by the Naïve, SES or the Holts Method. Whilst the OUT policy, Naïve, SES and Holt’s Method are all available native in many commercial software packages Damped Trend is less common and those that we know of do not allow for negative parameter values. We believe that the DT forecasting methodology deserves much more attention as a supply chain forecasting method in the OR / OM literature than it currently receives.

8. Appendix. Obtaining the stability via Jury's Inners

The general form of a transfer function is $TF(z)=\frac{B(z)}{A(z)}=\left(\sum_{i=0}^{m} b_i z^i\right)\Big/\left(\sum_{i=0}^{n} a_i z^i\right)$ where n is the order of the transfer function and $m \le n$. The denominator of a system transfer function is denoted $A(z)$ (to avoid confusion, here $A(z)$ is the denominator of (5)). Jury (1971) shows that the necessary and sufficient conditions for stability of a linear discrete system are given by: $A(1)>0$, $(-1)^n A(-1)>0$, and the matrices $\mathbf{\Delta}_{n-1}^{\pm}=\mathbf{X}_{n-1}\pm\mathbf{Y}_{n-1}$ are positive innerwise. The matrices

$$\mathbf{X}_{n-1}=\begin{bmatrix} a_n & a_{n-1} & a_{n-2} & \cdots & a_2 \\ 0 & a_n & a_{n-1} & \cdots & a_3 \\ 0 & 0 & a_n & \cdots & a_4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix},\ \mathbf{Y}_{n-1}=\begin{bmatrix} 0 & 0 & 0 & \cdots & a_0 \\ \vdots & \vdots & \vdots & ⋰ & \vdots \\ 0 & 0 & a_0 & \cdots & a_{n-4} \\ 0 & a_0 & a_1 & \cdots & a_{n-3} \\ a_0 & a_1 & a_2 & \cdots & a_{n-2} \end{bmatrix}, \tag{24}$$

are made up of the co-efficients a_i of z in $A(z)$. Then, $A(z)$ can be expressed as

$$A(z)=\underbrace{(1-\alpha)\phi}_{a_0}+z\,\underbrace{(\alpha-\phi-1+\alpha\beta\phi)}_{a_1}+z^2\underbrace{1}_{a_2}. \tag{25}$$

Taking each criterion in turn:

- $A(1)$ must be greater than zero: $A(1)=A(z)\big|_{z\to 1}$, that is $A(1)$ is given by (25) with the z is replaced with 1,

$$A(1)=\alpha\left(1+\phi(\beta-1)\right)>0 \tag{26}$$

- $(-1)^n A(-1)>0$ must be greater than zero. In the same manner as above, $(-1)^n A(-1)$ is given by (25) with z replaced by -1 and $n=2$,

$$(-1)^n A(-1)=\alpha-2+\phi\left(\alpha-2+\alpha\beta\right)>0 \tag{27}$$

- $\mathbf{\Delta}_{n-1}^{\pm}=\mathbf{X}_{n-1}\pm\mathbf{Y}_{n-1}$ must be positive innerwise. A matrix is positive innerwise if its determinant is positive and all the determinants of its Inners are also positive. Because (25) is only of second order, $n=2$, then the $\mathbf{\Delta}_{n-1}^{\pm}$ matrices only contain one element (Disney, 2008). Thus, to ensure that the elements of (24) are positive innerwise it is enough, in our system here, that

$$\left.\begin{aligned} \mathbf{\Delta}_{n-1}^{+} &= a_2+a_0 = 1+(1-\alpha)\phi>0 \\ \mathbf{\Delta}_{n-1}^{-} &= a_2-a_0 = 1-(1-\alpha)\phi>0 \end{aligned}\right\}. \tag{28}$$

9. References

Acar, Y., Gardner, E.S. Jr, 2012. Forecasting method selection in a global supply chain. *International Journal of Forecasting*, 28 (4), 842-8.

Adelson, R.M., 1966. The dynamic behaviour of linear forecasting and scheduling rules. *Operational Research Quarterly,* 17 (4), 447-62.

Axsäter, S., 2006. *Inventory Control*, Springer, New York, USA.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 2008. *Time Series Analysis: Forecasting and Control*. New Jersey: John Wiley & Sons.

Brenner, J.L., D'Esopo, D.A., Fowler, A.G., 1968. Difference equations in forecasting formulas. *Management Science*, 15 (3), 141-59.

Brown, R.G., 1963. *Smoothing, Forecasting and Prediction of Discrete Time Series*. New Jersey: Prentice-Hall.

Chen, F., Ryan, J.K., Simchi-Levi, D., 2000. The Impact of exponential smoothing forecasts on the bullwhip effect. *Naval Research Logistics*, 47 (4), 269-86.

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 (3), 567-90.

Disney, S.M., 2008. Supply chain aperiodicity, bullwhip and stability analysis with Jury's inners. *IMA Journal of Management Mathematics*, 19 (2), 101-16.

Disney, S.M., Gaalman, G., Hosoda, T., 2012. Review of stochastic cost functions for production and inventory control. In: Grubbström, R.W., Hinterhuber, H.H., eds. *Pre-prints of the 17th International Working Seminar of Production Economics*. Innsbruck, Austria. 20-24 February. Vol. 1, pp. 117-28.

Fildes, R., Nikolopous, K., Crone, S., Syntetos, A., 2008. Forecasting and operational research: a review. *Journal of the Operational Research Society*, 59 (9), 1150-72.

Gaalman, G., Disney, S.M., 2012. Bullwhip behavior in the Order-Up-To policy with ARIMA demand. *4th World Production and Operations Management Conference,* Amsterdam, The Netherlands, 1st-5th July.

Gardner, E.S. Jr, 1985. Exponential smoothing: the state of the art. *Journal of Forecasting*, 4 (1), 1-28.

Gardner, E.S. Jr, 1990. Evaluating forecast performance in an inventory control system. *Management Science*, 36 (4), 490-99.

Gardner, E.S. Jr, McKenzie, E., 1985. Forecasting trends in time series. *Management Science,* 31 (10), 1237-46.

Gardner, E.S. Jr, McKenzie, E., 1989. Seasonal exponential smoothing with damped trends. *Management Science*, 35 (3), 372-76.

Gardner, E.S. Jr, McKenzie, E., 2011. Why the damped trend works. *Journal of the Operational Research Society*, 62 (6), 1177-80.

Gilchrist, W.G., 1976. *Statistical Forecasting*. London: Wiley.

Hoberg, K., Thonemann, U.W., Bradley, J.R., 2007. Analyzing the effect of the inventory policy on order and inventory variability with linear control theory. *European Journal of Operational Research*, 176 (3), 1620-42.

Holt, C.C., 2004. Forecasting seasonals and trends by exponential weighted moving averages. *International Journal of Forecasting*, 20 (1), 5-10.

Hosoda, T., Disney, S.M., 2006. On variance amplification in a three-echelon supply chain with minimum mean squared error forecasting. *OMEGA: The International Journal of Management Science*, 34 (4), 344-58.

Hyndman, R.J., Akram, M., Archibald, B.C., 2008. The admissible parameter space for exponential smoothing methods. *Annals of the Institute of Statistical Mathematics,* 60 (2), 407-26.

Johnson, G.D., and Thompson H.E., 1975. Optimality of myopic inventory policies for certain dependent demand processes. *Management Science,* 21 (11), 1303-1307.

Jury, E.I., 1971. Inners approach to some problems of system theory. *IEEE Transactions on Automatic Control*, 16 (3), 233-40.

Lalwani, C.S., Disney, S.M., Towill, D.R., 2006. Controllable, observable and stable state space representations of a generalized Order-Up-To policy. *International Journal of Production Economics*, 101 (1), 173-84.

Makridakis, S., Hibon, M., 2000. The M3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16 (4), 451-76.

McClain, J.O., Thomas, L.J., 1973. Response-variance tradeoffs in adaptive forecasting. *Operations Research*, 21 (2), 554-68.

Nise, N.S., 2004. *Control systems Engineering* 4th ed. Hobeken, New Jersey: John Wiley & Sons.

Ouyang, Y., Daganzo, C., 2006. Characterization of the bullwhip effect in linear, time-invariant supply chains: Some formulae and tests. *Management Science*, 52 (10), 1544-56.

Popplewell, K., Bonney, M.C., 1987. The application of discrete linear control theory to the analysis and simulation of multi-product, multi-level production control systems. *International Journal of Production Research*, 25 (1), 45-56.

Roberts, S.A., 1982. A general class of Holt-Winters type forecasting models. *Management Science*, 28 (8), 808-20.

Snyder, R.D., Koehler, A.B., Ord, K., 2002. Forecasting for inventory control with exponential smoothing. *International Journal of Forecasting*, 18 (1), 5-18.

Tashman, L.J., Kruk, J.M., 1996. The use of protocols to select exponential smoothing procedures: a reconsideration of forecasting competitions. *International Journal of Forecasting*, 12 (2), 235-53.

Taylor, J.W., 2003. Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting*, 19 (4), 715-25.

Towill, D.R., 1982. Dynamic analysis of an inventory and order based production control system. *International Journal of Production Research*, 20 (6), 671-87.

Tsypkin, Y.Z., (1964). *Sampling systems theory and its application*. Pergamon Press, Oxford.

Vassian, H.J., 1955. Application of discrete variable servo theory to inventory control. *Journal of the Operations Research Society of America*, 3 (3), 272-82.

Veinott, A.F., 1965. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Management Science*, 12 (3), 206-222.

Wang, X., Disney, S.M., Wang, J., 2013. Exploring the oscillatory dynamics of a forbidden returns inventory system. Forthcoming in the *International Journal of Production Economics*. DOI: 10.1016/j.ijpe.2012.08.013.

Winters, P.R., 1960. Forecasting sales by weighted moving averages. *Management Science*, 6 (3), 324-42.