

Behavioral causes of the bullwhip effect: an analysis using linear control theory

Citation for published version (APA):

Udenio, M., Vatamidou, E., Fransoo, J. C., & Dellaert, N. P. (2017). Behavioral causes of the bullwhip effect: an analysis using linear control theory. *IIE Transactions*, 49(10), 980-1000.
<https://doi.org/10.1080/24725854.2017.1325026>

DOI:

[10.1080/24725854.2017.1325026](https://doi.org/10.1080/24725854.2017.1325026)

Document status and date:

Published: 03/10/2017

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Behavioral causes of the bullwhip effect: An analysis using linear control theory

Maximiliano Udenio, Eleni Vatamidou, Jan C. Fransoo & Nico Dellaert

To cite this article: Maximiliano Udenio, Eleni Vatamidou, Jan C. Fransoo & Nico Dellaert (2017) Behavioral causes of the bullwhip effect: An analysis using linear control theory, IISE Transactions, 49:10, 980-1000, DOI: [10.1080/24725854.2017.1325026](https://doi.org/10.1080/24725854.2017.1325026)

To link to this article: <http://dx.doi.org/10.1080/24725854.2017.1325026>



© 2017 The Author(s). Published with license by Taylor & Francis© Maximiliano Udenio, Eleni Vatamidou, Jan C. Fransoo, and Nico Dellaert



Accepted author version posted online: 04 May 2017.
Published online: 04 May 2017.



Submit your article to this journal [↗](#)



Article views: 516






View related articles [↗](#)



View Crossmark data [↗](#)

Behavioral causes of the bullwhip effect: An analysis using linear control theory

Maximiliano Udenio ^a, Eleni Vatamidou^b, Jan C. Fransoo ^a and Nico Dellaert ^a

^aSchool of Industrial Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands; ^bInria Centre de Recherche Sophia Antipolis Mediterranee, Provence-Alpes-Cote d'Azur, France

ABSTRACT

It has long been recognized that the bullwhip effect in real life depends on a behavioral component. However, non-experimental research typically considers only structural causes in its analysis. In this article, we study the impact of behavioral biases on the performance of inventory/production systems modeled through an APVIOBPCS (Automatic Pipeline, Variable Inventory, Order-Based Production Control System) design using linear control theory. To explicitly model managerial behavior, we allow independent adjustments to inventory and pipeline feedback loops. We consider the biases of smoothing/over-reaction to inventory and pipeline mismatches and the under-/over-estimation of the pipeline. To quantify the performance of the system, we first develop a new procedure to determine the exact stability region of the system and we derive an asymptotic stability region that is independent of the lead time. Afterwards, we analyze the effect of different demand signals on order and inventory variations. Our findings suggest that normative policy recommendations must take demand structure explicitly into account. Finally, through extensive numerical experiments, we find that the performance of the system depends on the combination of the behavioral biases and the structure of the demand stream.

ARTICLE HISTORY

Received 15 January 2015
Accepted 19 April 2017

KEYWORDS

Control theory; bullwhip effect; supply chain management; stability; behavioral operations


1. Introduction

The bullwhip effect is a major problem in today's supply chains; it is receiving considerable research attention. Lee *et al.* (1997b) define it as the observed propensity for material orders to be more variable than demand signals and for this variability to increase the further upstream a company is in a supply chain. It is a dynamic phenomenon that has sparked a vast body of research from a wide array of methodologies. Empirical, experimental, and analytic studies of the bullwhip exist of both a descriptive and a normative nature. The causes for the bullwhip effect can be broadly separated into operational (such as order batching and price fluctuations) and behavioral categories (such as artificially inflating orders and pipeline under-estimation). In this article, we use control theory as a modeling methodology to study behavioral causes of the bullwhip effect.

Regarding behavior, it has long been understood that decision makers do not operate under the paradigm of complete rationality (Su, 2008). Both in real-life and in experiments, humans operate in ways that deviate from theoretical predictions. We make mistakes. We exhibit psychological biases that affect our decisions. In operations research, heuristics containing feedback structures are commonly used to model human behavior: decisions bring about changes that affect future decisions. The modeling of feedback structures in a supply chain context can be traced back to Forrester (1958) and the introduction of system dynamics. Such modeling makes it possible to understand the dynamics of the system under study and

how they are affected by non-observable parameters and their interactions. In particular, anchor and adjustment heuristics (Tversky and Kahneman, 1974) are used to model the feedback loops introduced by decisions pertaining to the generation of material orders. In these heuristics, forecasts act as an anchor to orders and deviations from inventory and pipeline targets drive order adjustments up or down (Serman, 1989). Behavioral research on the bullwhip effect is primarily descriptive, linking deviations from optimal adjustments to human biases. Experimental work shows that people consistently under-estimate the pipeline inventory when making ordering decisions (Serman, 1989; Rong *et al.*, 2008; Croson *et al.*, 2014). These biases, linked to the appearance of the bullwhip effect, were also observed in empirical data (Udenio *et al.*, 2015).

On the analytical front, however, these behavioral biases and their connection to the bullwhip effect have not been explicitly studied. In this article, we do so by investigating the influence of a number of behavioral biases on the performance of Automatic Pipeline, Variable Inventory, Order Based Production Systems (APVIOBPCS). We offer three distinct contributions to the literature. First, we derive a series of exact conditions for the stability of a general APVIOBPCS design with arbitrary lead time and find an asymptotic region for the stability of the system for all lead times. Second, we perform an extensive numerical study where we explicitly model behavioral biases through the independent adjustment of inventory and pipeline feedback loops. In particular, we consider the biases of smoothing/over-reaction to inventory and pipeline mismatches and the

CONTACT Maximiliano Udenio  m.udenio@tue.nl

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uiie.

©2017 Maximiliano Udenio, Eleni Vatamidou, Jan C. Fransoo, and Nico Dellaert. Published with license by Taylor & Francis.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.

under/over-estimation of the pipeline requirements. We investigate different performance dimensions (i.e., stationary bullwhip, worst-case amplification, and integral time-weighted absolute error) and analyze their tradeoffs. Third, we present managerial insights linking the results of our study to their potential application on behavioral research as well as industrial practice. We highlight the influence of the structure of demand on the impact of behavior and the ensuing consequences that this has with regards to the implementation of specific results in practical applications.

A common way to investigate feedback-ridden dynamic inventory models is with the aid of linear control theory. Its use in inventory models traces back to Simon (1952), who studied a continuous-time system through the Laplace transform method, in which inventory targets are used to derive material orders. Vassian (1955) studied the equivalent discrete-time system using the Z-transform, and Deziel and Eilon (1967) extended the discrete case by adding a smoothing parameter to control the variance of the response. Towill (1982) extended and formalized these ideas with the introduction of the Inventory and Order-Based Production Control System (IOBPCS) design framework, the predecessor of the models studied in this article. In an IOBPCS design, replenishment orders are generated as the sum of an exponentially smoothed demand forecast and a fraction of the inventory discrepancy (the gap between a constant target inventory and the actual value), which acts as a feedback loop. His work, by representing an inventory/production system in block diagram form, allowed for the straightforward application of linear control theory methodologies to study its structural and dynamic properties.

A first extension to IOBPCS is VIOBPCS (Variable Inventory and Order-Based Production Control System), where the inventory target is no longer constant but rather is calculated each period as a multiple of the demand forecast. Edghill and Towill (1990) studied this system and found that, in comparison with IOBPCS, the variable inventory targets of VIOBPCS designs introduce interesting tradeoffs between the “marketing” and “production” sides of a firm and increased service levels through a better correlation of inventory and demand, at the cost of increased variability in orders. A powerful extension, APIOBPCS (Automatic Pipeline Inventory and Order Based Production Control System), adds a second feedback loop in the form of a pipeline adjustment (John *et al.*, 1994). The ordering logic of this design is a direct equivalent to the anchor and adjustment heuristic commonly used to model beer-game playing behavior (Sternan, 1989; Croson *et al.*, 2014).

Discrete APIOBPCS and its Variable Inventory extension, APVIOBPCS have been extensively studied through the Z-transform method in the two decades since the work of John *et al.* (1994). Dejonckheere *et al.* (2003) showed that APVIOBPCS designs with full adjustments (inventory and pipeline discrepancies are filled every period) are equivalent to Order-Up-To (OUT) policies. Such policies are often implemented in practice with inflated lead times used to represent the safety stock (i.e., a safety lead time). Hoberg *et al.* (2007a) used this design (which we denote as APVIOBPCS-OUT) to compare the stationary and transient responses of echelon and installation stock policies in a two-echelon supply chain. Hoberg *et al.* (2007b) used the same system design but focused instead on quantifying the effect of the forecasting smoothing

parameter α . They found that both echelon stock policies and values of α close to zero contribute to a reduction of inventory and order amplification. In a recent study, Hoberg and Thonemann (2014) extended the above setup to allow for information delays in addition to material delays. They found that echelon stock policies are severely hampered by such delays and thus proposed a compensated echelon policy to counteract this effect.

An extension to the APVIOBPCS-OUT design is the general APVIOBPCS design, where the inventory and pipeline feedback loops can be adjusted to fine-tune the response of the system, thus controlling the desired speed to adjust the inventory and pipeline toward their targets. Disney and Towill (2006) studied a particular subset of parameter settings: DE-APVIOBPCS designs, christened after the work of Deziel and Eilon. In these designs, the adjustments for inventory and pipeline are always equal and fractional, which is equivalent to a Generalized-OUT policy with smoothing of the inventory position adjustment. This constraint in the model parameters has very attractive properties. From an optimal design perspective, DE-designs always produce stable and aperiodic responses; from a mathematical perspective, DE-designs produce tractable, elegant expressions. General APVIOBPCS designs with independent parameters for inventory and pipeline adjustments, on the contrary, enable the modeling of a larger number of policies but exhibit none of these desirable characteristics. The value of the feedback controller parameters determines whether or not the system is stable. The complexity introduced by independent parameters is such that a substantial number of studies is dedicated to characterizing the stability of APVIOBPCS designs. Disney and Towill (2002) found a general expression for stability boundaries by modeling lead times as a third-order lag instead of a pure delay. Warburton *et al.* (2004) defined stability boundaries for an equivalent continuous-time system by finding the exact solution to its characteristic equation. Disney, Towill, and Warburton (2006) derived a stability criterion through a continuous, time domain, differential equation approach. The latter also showed that, although quantitatively different, the various continuous lead time approximations share many qualitative interpretations with their discrete-time counterparts. In the discrete domain, Disney (2008) demonstrated the use of “Jury’s Inners” (Jury, 1964, p. 85) to derive stability conditions for a given lead time and, finally, Wei *et al.* (2013) found conditions for stability through the analysis of the difference equation representation of the system.

In addition to the above discussion on stability, there is a large body of literature dedicated to the analysis of other aspects of the APVIOBPCS design and its variants. Table 1 shows a summary of the main research objectives and design setup of the papers closest to this study. For a more comprehensive literature review on the application of control theory to inventory/production systems, we refer the reader to Ortega and Lin (2004) and Sarimveis *et al.* (2008). The former centers in the application of classic control, whereas the latter pays special attention to advanced control methodologies. For a different perspective on the application of APVIOBPCS designs, Zhou *et al.* (2010) reviewed the literature considering the “pragmatic role” of these designs and how they can aid in the decision-making process for different production situations. In a recent study, Disney *et al.* (2013) described how Lexmark implemented an APVIOBPCS design into their

Table 1. Literature on APVIOBPCS designs and main extensions.

Paper	Inventory Policy	Demand	Forecast	Lead time	Parameters	Objective
Hosoda and Disney (2006)	APVIOBPCS-OUT	AR(1)	MMSE forecast	$L_{SC} = 15$	—	Analysis of BW for a three-echelon supply chain
Hoberg et al. (2007b)	APVIOBPCS-OUT	Stationary i.i.d., step change, Sinusoidal	Exponential smoothing	$L = 4$	α	Analysis of stationary performance of echelon and installation-stock policies in two-stage supply chains
Hoberg et al. (2007a)	APVIOBPCS-OUT	Stationary i.i.d., step change, Sinusoidal	Exponential smoothing	$L = 4$	α	Analysis of non-stationary performance of echelon and installation-stock policies in two-stage supply chains
Hoberg and Thonemann (2014)	APVIOBPCS-OUT	Stationary i.i.d., step change, Sinusoidal	Exponential smoothing	$L = 2, 4$	α, τ	Analysis of stationary and non-stationary performance of echelon and installation-stock policies in two-stage supply chains with information delay τ
Li et al. (2014)	APVIOBPCS-OUT	Sinusoidal, real-life pattern	Naive, exponential smoothing, Holts, dampened trend	$L = 1$	α	Analysis of the bullwhip response of an OUT policy under different forecasting methodologies
Disney et al. (2008)	DE-APVIOBPCS	Stationary i.i.d.	Exponential smoothing	$L = 1$	α, γ_I	Exploration of four different strategies for coordination in a two-Echelon Supply Chain
Disney, Towill, and Warburton (2006)	DE-APVIOBPCS	Stationary i.i.d.	Exponential smoothing	$L = 1, \dots, 5$	α, γ_I	Guideline for parameter setting for robust bullwhip (BW) reduction
Disney, Farasyn, and Lambrecht (2006)	DE-APVIOBPCS	i.i.d., AR(1), MA(1), and ARMA(1,1)	Mean demand, exponential smoothing	$L = 2$	α, γ_I	Parameter tuning to exploit supply chain characteristics (min inventory exposure, min variability)
Disney and Grubbström (2004)	DE-APVIOBPCS	AR(1)	Exponential smoothing	$L = 1$	α, γ_I	Analysis of economic performance considering holding, backlog, and production costs
Disney et al. (2004)	DE-APVIOBPCS	Stationary i.i.d.	Exponential smoothing, mean demand	$L = 3$	α, γ_I	Closed form calculation of order and inventory variance. Optimal γ_I for weighted sum of variances
Disney and Towill (2003)	DE-APVIOBPCS	Stationary i.i.d., step change	Exponential smoothing	$L = 3$	α, γ_I	Development of stationary (weighted sum of order and inv BW) and dynamic (inv integrated time-weighted average error) performance metrics
Hoberg and Thonemann (2015)	DE-APVIOBPCS	Stationary i.i.d., step change, Sinusoidal	Exponential smoothing	$L = 4$	α, γ_I	Analysis of stationary and dynamic performance from a total cost perspective
Dejonckheere et al. (2003)	DE-APVIOBPCS, APVIOBPCS	Stationary i.i.d., step change, Sinusoidal	Exponential smoothing, moving average, demand signal processing (DSP)	$L = 3$	$\alpha, \gamma_P, \gamma_I$	Comparison of BW generated by four different policies. Use of Frequency Response plots to infer response characteristics to correlated demands
Dejonckheere et al. (2004)	DE-APVIOBPCS, APVIOBPCS	Stationary i.i.d. (information sharing)	Exponential smoothing, moving average, DSP	$L = 3$	$\alpha, \gamma_P, \gamma_I$	Comparison of BW on a four-echelon Supply Chain with different policies with and without information sharing
Disney, Towill, and Warburton (2006)	DE-APVIOBPCS, APVIOBPCS	Stationary i.i.d., step change	Exponential smoothing	$L = 1, \dots, 3$	$\alpha, \gamma_P, \gamma_I$	Demonstration of the equivalence of control theoretic, differential, and difference equation solution approaches w.r.t. managerial insights
Disney (2008)	DE-APVIOBPCS, APVIOBPCS	Stationary i.i.d., unit impulse	Mean demand	$L = 2$	$\alpha, \gamma_P, \gamma_I$	Demonstrate the use of Jury's inners to calculate stability conditions. Aperiodicity via pole analysis
Zhou et al. (2010)	DE-APVIOBPCS, APVIOBPCS	Stationary i.i.d., step change	Exponential smoothing	$L = 4, 8, 12$	$\alpha, \gamma_P, \gamma_I$	Guideline for parameter setting for BW reduction based on pragmatic company goals

Note: We represent all studies with the equivalent notation used in this article.

planning activities to eliminate the bullwhip from their toner operations.

In terms of model assumptions, we allow for independent pipeline and inventory adjustments, consistent with Dejonckheere *et al.* (2004) and Disney and Towill (2006). In terms of performance metrics, we adopt the stationary and dynamic measures of Hoberg *et al.* (2007b) and Hoberg and Thonemann (2015). Conceptually, however, our objective is to relate parameter combinations to different behavioral biases. In particular, we use the inventory and pipeline feedback loops to explicitly model smoothing and over-reaction to inventory and pipeline mismatches and the under- and over-estimation of the pipeline. Hence, our experimental setup is informed by the behavioral operations literature, in particular beer-game-based bullwhip effect research (Sternan, 1989; Croson *et al.*, 2014). We find a mixed picture regarding the impact of behavioral biases. Behavioral biases do not consistently deteriorate performance; their influence depends on the structure of the demand stream. Thus, a behavior that increases the performance of the system under independent and identically distributed (i.i.d.) normal demand (e.g., order smoothing) can be detrimental in the presence of a shock. This implies that a comprehensive analysis of the influence of behavior on the bullwhip effect must take demand assumptions explicitly into account.

The same also holds true for the structural causes of the bullwhip. Demand forecast updating is a known structural cause of the bullwhip (Lee *et al.*, 1997a; Dejonckheere *et al.*, 2003), and a substantial body of literature relates specific forecasting methodologies to the bullwhip effect. Authors consistently report that, whereas forecasting itself induces it, the measured bullwhip depends on the combination of forecasting methodology, ordering policy, and demand distribution; see, for example, the analytical work of Chen, Drezner, Ryan, and Simchi-Levi (2000) and Zhang (2004); system dynamics simulations by Wright and Yuan (2008); and control-theoretic analysis of the topic by Dejonckheere *et al.* (2002). With regards to the specific forecasting methodologies adopted in this article, we refer the reader to Chen, Ryan, and Simchi-Levi (2000) and Hoberg *et al.* (2007b) for explicit analyses of the effect of the exponential smoothing parameter α on the magnitude of the bullwhip effect under APVIOBPCS-OUT policies and Hoberg and Thonemann (2015) for the effect of α on DE-APVIOBPCS policies. We analyze the influence of α on general APVIOBPCS policies in a series of numerical experiments presented in Appendix A.

The rest of this article is structured as follows. In the next section, we introduce our discrete-time model and the performance metrics. We continue, in Section 3, with a comprehensive analysis of the stability of the system, deriving exact expressions for the general stability boundaries. In Section 4, we analyze the stationary and dynamic responses of the system as a function of the behavioral biases and identify the tradeoffs among them. We provide our conclusions in Section 5. Finally, we present additional numerical experiments and all mathematical proofs in Appendix B.

2. Model description and performance metrics

In this section, we analyze a discrete-time, periodic-review, single-echelon, general APVIOBPCS design with an

exponentially smoothed forecast of demand. The structural parameters of the system are the inventory coverage ($C \in \mathbb{R}^+$), the delivery lead time ($L \in \mathbb{N}$), and the forecast smoothing parameter $\alpha \in [0, 2]$. The system maintains a target inventory (\hat{i}) equal to the expected demand over C periods and a target pipeline (\hat{p}) equal to the expected lead time demand. The lead time is assumed deterministic and defined as the time elapsed between the placement and receipt of a replenishment order. Replenishment orders (o) depend on the actual values of the inventory (i) and pipeline (p). The behavioral parameters of the system are the inventory ($\gamma_I \in \mathbb{R}$) and pipeline ($\gamma_P \in \mathbb{R}$) adjustment factors. The behavioral parameters specify the fraction of the gap between target and actual values that are taken into account to calculate orders: γ_I is the fraction of the inventory gap to be closed, and γ_P is the fraction of the pipeline gap to be closed. For instance, a system with $\gamma_I = 1$ and $\gamma_P = 0$ completely closes the inventory gap with every order, whereas it ignores the pipeline entirely.

Formally, the sequence of events and the equations in the model are as follows: at the beginning of each period (t), a replenishment order (o_t) based on the previous period's demand forecast (f_{t-1}) is placed with the supplier. Following this, the orders that were placed L periods prior are received. Next, the demand for the period (d_t) is observed and served. Excess demand is back-ordered. Then, the demand forecast is updated according to the formula $f_t = \alpha d_t + (1 - \alpha) f_{t-1}$. The forecast represents the expected demand and is used to compute the target levels of both inventory, $\hat{i}_t = C f_t$, and pipeline, $\hat{p}_t = L f_t$. In such a policy, the target inventory level plays a role analogous to the safety stock in OUT policies (Dejonckheere *et al.*, 2003). This concept is widely applied in practice, where C can be computed to hedge against lead time and forecast variability to satisfy any arbitrary service level (Hoberg and Thonemann, 2014). For example, we can compute C using the traditional safety stock definition based on the standard deviation of the forecast error; i.e., setting $C\hat{d} = k\hat{\sigma}_t^L$, where \hat{d} is an estimate of the mean demand, k is a constant chosen to meet a desired service level, and $\hat{\sigma}_t^L$ an estimate of the standard deviation of the L period forecast error (see Eppen and Martin (1988) and Chen, Ryan, and Simchi-Levi (2000)) for approximate and exact methods to determine this parameter). Note that the target inventory defined by this policy is equal in expectation to the safety stock computed through the traditional calculation. However, at any given period, the target inventory computed by the APVIOBPCS design will under-estimate (over-estimate) the traditional safety stock when the period's forecast is smaller (larger) than the estimate of the mean demand. The magnitude of the under-/over-estimation depends on the ratio f_t/\hat{d} and, thus, on the coefficient of variation of the demand and the smoothing parameter of the forecast. We refer the reader to Disney, Farasyn, and Lambrecht (2006) for an in-depth study regarding the customer service implications of the APVIOBPCS design.

The orders that will be placed in the following period (o_{t+1}) are generated according to an anchor and adjustment-type procedure, $o_{t+1} = \gamma_I(\hat{i}_t - i_t) + \gamma_P(\hat{p}_t - p_t) + f_t$. The balance equations for inventory (i) and pipeline (p) are $i_t = i_{t-1} + o_{t-1} - d_t$ and $p_t = p_{t-1} + o_t - o_{t-1}$. Note that the assumptions that orders and inventories can be negative are necessary to maintain the linearity of the model.

This sequence of events is identical to the one described in Hoberg *et al.* (2007a) with the difference being that our study introduces the fractional behavioral parameters γ_I and γ_P . Other studies of APVIOBPCS designs use a different order of events; e.g., in Dejonckheere *et al.* (2003) and Disney (2008), orders are placed at the end of each period. Such differences in the sequence of events introduce extra unit delays in the equations. However, these differences only affect the mathematical representation of the system; the structure of the system and the results remain the same. The model we introduce completely describes the relationships between the parameters of a general APVIOBPCS design. However, due to time dependencies, we cannot find a clear relationship between the inputs and the outputs of the system. For this reason, we turn from the time domain to the frequency domain (where these relationships become simply algebraic), by taking the Z-transform of the system's set of equations. The Z-transform is defined as

$$\mathcal{Z}\{x_t\} = X(z) = \sum_{t=0}^{\infty} x_t z^{-t}, \quad (1)$$

where z is a complex variable and x_t is the value of a time series at time t . We refer the reader to Jury (1964) and Nise (2011) for a comprehensive background on discrete systems and the Z-transform method and to Dejonckheere *et al.* (2003) and Hoberg *et al.* (2007a) for an introduction to their application on inventory modeling.

Using the following properties of the Z-transform:

$$\mathcal{Z}\{a_1 x_t + a_2 y_t\} = a_1 X(z) + a_2 Y(z) \quad (\text{linearity}), \quad (2)$$

$$\mathcal{Z}\{x_{t-T}\} = z^{-T} X(z) \quad (\text{time delay}), \quad (3)$$

we can write all system parameters in the frequency domain. The equation for orders (see two paragraphs back in this section) is re-written in the frequency domain as

$$O(z) = \frac{\gamma_I(\hat{I}(z) - I(z)) + \gamma_P(\hat{P}(z) - P(z)) + F(z)}{z}. \quad (4)$$

In control theory, the response of a system is completely characterized by its transfer function $G(z) = N(z)/C(z)$, which represents the change in output $N(z)$ with regards to a change in input $C(z)$ in the frequency domain. In this article, we are interested in studying the response of orders $O(z)$ and inventories $I(z)$ to changes in customer demand $D(z)$. Therefore, we are interested in the transfer function of orders ($G_O(z) = O(z)/D(z)$) and the transfer function of inventories ($G_I(z) = I(z)/D(z)$). We can then write the transfer function of orders in terms of the system parameters as

$$\begin{aligned} G_O(z) &= \frac{O(z)}{D(z)} = \frac{[\gamma_I(\hat{I}(z) - I(z)) + \gamma_P(\hat{P}(z) - P(z)) + F(z)] \frac{1}{z}}{D(z)} \\ &= \frac{[\alpha(\gamma_I C + \gamma_P L + 1)(z - 1) + \gamma_I(z - 1 + \alpha)] z^L}{(z - 1 + \alpha)(z^L(z - 1 + \gamma_P) + (\gamma_I - \gamma_P))}. \end{aligned} \quad (5)$$

Analogously, we write the transfer function of inventories:

$$\begin{aligned} G_I(z) &= \frac{I(z)}{D(z)} = \frac{\frac{z}{z-1}[O(z)z^{-L} - D(z)]}{D(z)} \\ &= \frac{z\alpha(\gamma_I C + \gamma_P L + 1)(z - 1) + z(z - 1 + \alpha)(\gamma_P - z^L(z - 1 + \gamma_P))}{(z - 1)(z - 1 + \alpha)(z^L(z - 1 + \gamma_P) + (\gamma_I - \gamma_P))}. \end{aligned} \quad (6)$$

In the following sections, we use the above transfer functions to analyze the stability and calculate the performance metrics for the system under consideration.

2.1. Stationary performance metrics

In this section, we use concepts from control theory to derive stationary performance metrics for an APVIOBPCS design through the lens of order and inventory variability. Order and inventory variability are intimately linked to the notion of the bullwhip effect—i.e., the propensity of orders to be more variable than demand signals—and for this variability to increase the further upstream a firm is in a supply chain (Lee *et al.*, 1997a). We define three different measures to quantify the amplification of variability according to the characteristics of the demand: the amplification ratio, the bullwhip measure, and the worst-case amplification. The amplification ratio describes the order and inventory behavior of the system as a response to a stationary demand of an arbitrary frequency. The bullwhip measure quantifies the behavior as a response to a normally distributed demand. The worst-case amplification quantifies the maximum amplification ratio of the system as a response to any possible demand.

2.1.1. Amplification ratio

A sinusoidal input to a linear system produces a sinusoidal output of the same frequency but of a different magnitude and phase. For a given linear system, the ratio between the amplitudes of the input and output at a given frequency is constant and is calculated as the modulus of its transfer function evaluated at that frequency (Dejonckheere *et al.*, 2003). Thus, for any input sinusoid, the steady-state amplification ratio of an APVIOBPCS design can be calculated directly from its transfer functions. Furthermore, it can be shown that for sinusoidal inputs, *the amplification ratio value is exactly the same as the ratio of the standard deviations of input over output* (Jakšič and Rusjan, 2008). Formally, for our system, we define $A_{O,\omega}$ as the amplification ratio of orders for a sinusoidal demand of frequency ω and $A_{I,\omega}$ as the amplification ratio of inventory for a sinusoidal demand of frequency ω , where

$$A_{O,\omega} = |G_O(e^{i\omega})|, \quad \text{and} \quad (7)$$

$$A_{I,\omega} = |G_I(e^{i\omega})|, \quad (8)$$

where $|G(e^{i\omega})|$ is the modulus of the transfer function evaluated at the frequency ω . It is important to note that our interest in the steady-state performance is not restricted to the expectation of a sinusoidal demand. Since any demand stream can be decomposed into a sum of sinusoids, analyzing the relevant frequency response plots (i.e., the graphical representation of the amplification ratio as a function of the demand harmonics with frequencies between zero and π) provides intuition about the performance of a system with regards to any arbitrary demand

pattern based on the amplitude of its constituent harmonics (Dejonckheere *et al.*, 2003).

2.1.2. The bullwhip measure

Disney and Towill (2003) defined the bullwhip measure as the ratio between input variance to output variance ($BW = \sigma_{out}^2 / \sigma_{in}^2$) and showed that if the mean of the input is zero and its variance is unity, the bullwhip of a system is proportional to the “noise bandwidth” metric commonly used in communications engineering. This measure has the intuitive representation of the square of the area below the frequency response plots. If the input to our system is a stationary i.i.d. normal demand stream, then the bullwhip of orders (BW_O) can be calculated through

$$BW_O = \frac{\sigma_O^2}{\sigma_D^2} = \frac{1}{\pi} \int_0^\pi |G_O(e^{i\omega})|^2 d\omega, \quad (9)$$

and the bullwhip of inventories (BW_I) is defined analogously:

$$BW_I = \frac{\sigma_I^2}{\sigma_D^2} = \frac{1}{\pi} \int_0^\pi |G_I(e^{i\omega})|^2 d\omega. \quad (10)$$

The assumption of an i.i.d. normal input ensures that all frequencies are equally represented in any given demand stream. Ouyang and Daganzo (2006) argued that, in a multi-echelon context, this assumption is too restrictive and results in a measure that is not robust. They showed that the only way to predict the appearance of the bullwhip effect under arbitrary multi-echelon chains is to study the worst-case amplification ratio of the system.

2.1.3. Worst-case amplification

Insights from the single-echelon bullwhip measures defined above cannot be extrapolated to the multi-echelon case. If the amplification ratio for any given frequency is larger than one, then a supply chain of identical echelons will always result in a bullwhip N stages upstream, even if the bullwhip measure is smaller than one and the demand is normally distributed (the number of stages N required to see this effect depends on the particular policy and demand stream). To make up for this limitation, researchers have proposed the use of the worst-case amplification as a complementary performance metric (Ouyang and Daganzo, 2006; Hoberg *et al.*, 2007b). Intuitively, the worst-case amplification corresponds to the maximum amplification ratio across all frequency components $\omega \in [0, \pi)$. We denote the worst-case amplification for orders as A_O^∞ and the worst-case amplification for inventories as A_I^∞ . We define them formally as

$$A_O^\infty = \sup_{\forall \omega \in [0, \pi)} |G_O(e^{i\omega})| \quad (11)$$

and

$$A_I^\infty = \sup_{\forall \omega \in [0, \pi)} |G_I(e^{i\omega})|. \quad (12)$$

This performance metric is robust. In other words, under any arbitrary demand, a supply chain constructed from N systems each with $A_O^\infty < 1$ ($A_I^\infty < 1$) will not amplify orders (inventories).

2.2. Dynamic performance

To study the dynamic performance of the system, we consider its behavior in the time domain after experiencing a shock. We quantify the system’s dynamic performance through the Integral Time-Weighted Absolute Error (ITAE), a measure of the performance of the system in terms of time-weighted deviations from the ideal response (Hoberg *et al.*, 2007a). We introduce a shock in the system in the form of a one-time step change in demand. Formally, the ITAE is defined as

$$ITAE = \sum_{t=0}^{\infty} t |\epsilon_t|, \quad (13)$$

where ϵ_t represents the absolute error between the actual response at time t and the steady-state response. This measure penalizes deviations from the new (target) steady state and introduces a linear penalty for longer-lasting deviations. Thus, both the amplification (i.e., how large the error is) and the settling time (i.e., how long it takes for the actual response to converge to the steady state response) of the system play a role in its determination. For our system, we define the $ITAE_O$ for orders and $ITAE_I$ for inventories. Formally,

$$ITAE_O = \sum_{t=0}^{\infty} t |o_t - d_t|, \quad (14)$$

$$ITAE_I = \sum_{t=0}^{\infty} t |i_t - (Cd_t)|. \quad (15)$$

Taken together, the different (stationary and dynamic) metrics discussed above give an overall impression of the performance of a particular system under a large number of demand assumptions. However, before studying the performance of any given system, we must define a necessary condition for its feasibility/stability. In Section 3, we derive analytic conditions for the stability of the system as a function of structural and behavioral parameters. Then, in Section 4, we use the stationary and dynamic metrics to study the influence of different behavioral biases on the overall performance of an APVIOBPCS system.

3. Stability and aperiodicity

A stable dynamic system yields a finite output for any finite input. In our model, the customer demand is the input, and orders and inventory are outputs. Hence, in this context, stability guarantees finite orders and inventories as a response to changes in demand—a pre-condition for any real-life system. We have the following formal definition for the stability of a system.

Definition 1. (Nise, 2011, p. 302). A system is stable if every bounded input yields a bounded output and unstable if at least one bounded input yields an unbounded output.

Although Definition 1 formally describes the stability of the system, it does not specify mathematical conditions necessary to test whether a given system is stable or not. An alternative definition-condition that connects the stability of a system with its transfer function is the following.

Definition 2. (Jury, 1964, p. 80). Suppose that $G(z) = N(z)/C(z)$ is the transfer function of a linear, time-invariant system and that the denominator $C(z)$ has exactly n roots p_i , namely, $C(p_i) = 0$, $i = 1, \dots, n$. We call the roots p_i poles of the transfer function, and we say that a system is stable if all poles p_i are within the unit circle of the complex plane ($|p_i| < 1$), marginally stable if at least one pole is on the unit circle ($|p_i| = 1$), and unstable if at least one pole resides outside the unit circle ($|p_i| > 1$).

Consequently, judging the stability of a system is equivalent to finding the solutions to the characteristic equation $C(z) = 0$.

Remark 1. Suppose that P with $|P| \geq 1$ is a root of $C(z)$ with multiplicity m , namely, $C^{(i)}(P) = 0$, $\forall i = 0, \dots, m-1$. If $N^{(i)}(P) = 0$, $\forall i = 0, \dots, m-1$, and if all other roots of $C(z)$ are inside the unit circle, then the system is called stabilizable. However, this is sometimes used alternatively as a definition for a stable system (Wunsch, 2005, p. 482). This is not the case here.

In the next section, following Definition 2, we derive sufficient conditions for the stability of the system through an analysis of the structure of the involved characteristic polynomials. We then introduce the aperiodicity of the system, a characterization of the dynamic response of a stable system. We begin our analysis with the response of orders to changes in demand, Equation (5), and follow with the analysis of the inventory response to changes in demand, Equation (6).

3.1. Stability boundaries

By comparing Equations (5) and (6), we see that the characteristic polynomials of orders and inventories are almost equal except for the extra factor $(z-1)$ that appears in the latter. The pole $z=1$ would render the inventory response marginally unstable, unless this is also a root of the numerator of $G_I(z)$ (see Equation (6) and Remark 1). To this effect, we use the geometric series identity $z^{L+1} - 1 = (z-1) \sum_{i=0}^L z^i$ to rewrite Equation (6) as

$$G_I(z) = \frac{z\alpha(\gamma_I C + \gamma_P L + 1) - z^{L+2} - z^{L+1}(\alpha + \gamma_P - 1) + \gamma_P z + \alpha\gamma_P(1 - \sum_{i=0}^L z^i)}{(z-1 + \alpha)(z^L(z-1 + \gamma_P) + (\gamma_I - \gamma_P))}. \quad (16)$$

Thus, in an APVIOBPCS design, the stability of both orders and inventories is defined by the same characteristic polynomial:

$$C(z) = (z-1 + \alpha)(z^L(z-1 + \gamma_P) + (\gamma_I - \gamma_P)). \quad (17)$$

Being a polynomial in z of degree $L+2$ with real coefficients, $C(z)$ has exactly $L+2$ roots. This polynomial is transcendental: it is impossible to find its roots independently of L . Furthermore, exact solutions for $C(z) = 0$ can only be found for values of $L \leq 2$. Thus, we study structural properties of $C(z)$ to derive a set of sufficient conditions that define an exact stability region for the general APVIOBPCS design. The proofs of all theorems, lemmas, and propositions are found in Appendix B.

It can be shown that APIOBPCS and APVIOBPCS policies share the same characteristic polynomial (Disney and Towill, 2006). Thus, the insights and conclusions derived from the analysis of $C(z)$ hold for APVIOBPCS designs. Therefore, we formulate our results for APVIOBPCS systems.

Proposition 1. *The stability of a general APVIOBPCS system with smoothing parameter $\alpha \in [0, 2)$ can be determined by analyzing the poles of the reduced characteristic polynomial:*

$$\hat{C}(z) := z^L(z-1 + \gamma_P) + (\gamma_I - \gamma_P). \quad (18)$$

An APVIOBPCS is stable if all the roots of $\hat{C}(z)$ are located inside the unit circle.

From the above proposition, we deduce that the stability of a general APVIOBPCS system with the commonly used exponential smoothing parameter range of $[0, 2)$ is completely determined by the values of L , γ_I , and γ_P . For $\alpha = 2$, $z = -1$ is a root of the characteristic polynomial $C(z)$ (see Equation (17)), which means that the system will be marginally stable unless $z = -1$ is also a root of the numerator of the transfer function. Therefore, if $\alpha = 2$, Proposition 1 holds when $\gamma_I C + \gamma_P L = 1$. In the following theorem, we specify the stability region of a general APVIOBPCS system in terms of a set of boundary conditions.

Theorem 1. *For each value of L , stability is guaranteed when γ_I and γ_P satisfy the following $L+1$ conditions:*

$$(i) \quad |\gamma_I - \gamma_P| < 1, \quad (19)$$

$$(ii) \quad (1 - (\gamma_I - \gamma_P)^2) |\gamma_P - 1|^{(n-1)} U_{n-1}(X) - |\gamma_P - 1|^n U_{n-2}(X) > 0, \quad n = 2, \dots, L, \quad (20)$$

where $U_n(X)$ is the Chebyshev polynomial of the second kind, defined by

$$U_n(X) = \frac{(X + \sqrt{X^2 - 1})^{n+1} - (X - \sqrt{X^2 - 1})^{n+1}}{2\sqrt{X^2 - 1}}, \quad (21)$$

with

$$X = \frac{1 - (\gamma_I - \gamma_P)^2 + (\gamma_P - 1)^2}{2|\gamma_P - 1|}, \quad (22)$$

and

$$(iii) \quad ((1 - (\gamma_I - \gamma_P)^2)^2 - ((\gamma_I - \gamma_P)(\gamma_P - 1))^2) \times |\gamma_P - 1|^{L-1} U_{L-1}(X) - 2(1 - (\gamma_I - \gamma_P)^2) |\gamma_P - 1|^L U_{L-2}(X) + |\gamma_P - 1|^{L+1} U_{L-3}(X) + 2(-1)^{L+1} (\gamma_I - \gamma_P)(\gamma_P - 1)^{L+1} > 0. \quad (23)$$

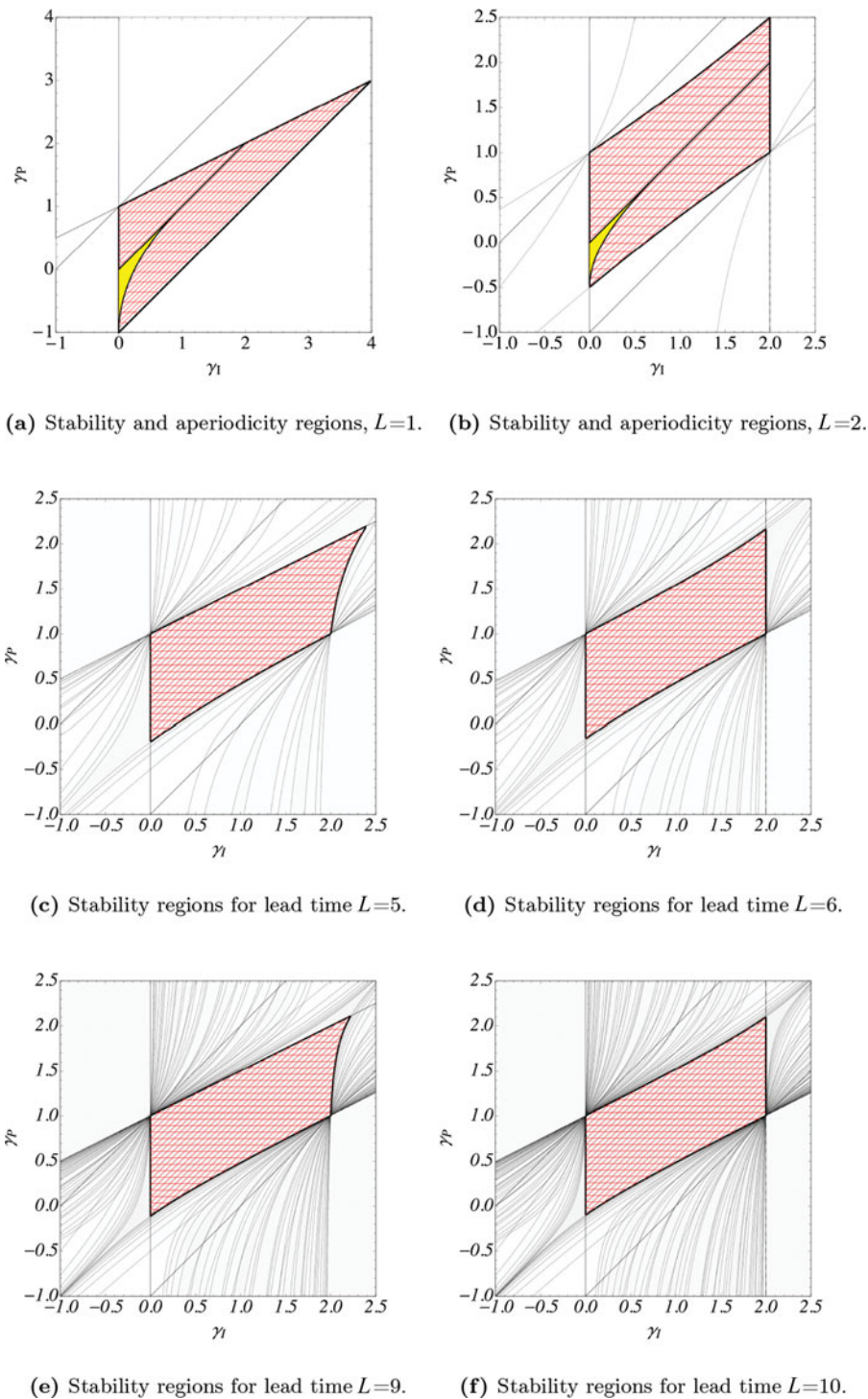


Figure 1. Stability and aperiodicity conditions.

Figure 1 shows a plot of all conditions defined by Theorem 1 for a number of different lead times. The three plots on the left correspond to odd lead times and the three plots on the right correspond to even lead times. The boundaries of each condition defined by Equations (19) to (23) are indicated by thin black lines. The area that satisfies all conditions simultaneously is clearly demarcated by a thick black line and a (red) mesh. The additional areas (yellow) in Figures 1(a) and 1(b) concern aperiodicity, which we define in Section 3.2. Note that the axes in Figure 1(a) are extended to show the entire region of stability.

Remark 2. It can be seen that the L conditions defined by Equations (19) and (20) describe regions of convergence decreasing in L . We observe that the intersection of all regions that are defined by the conditions in Equation (20) is equal to the region that is defined in Equation (20) for $n = L$. Moreover, the last condition can be simplified, as it produces the exact same region for all even (similarly for odd) lead times. This allows us to pose the following conjecture.

Conjecture 1. For each value of L , stability is guaranteed when γ_I and γ_P satisfy the following conditions:

- (i) $|\gamma_I - \gamma_P| < 1$.
- (ii) $(1 - (\gamma_I - \gamma_P)^2) |\gamma_P - 1|^{(L-1)} U_{L-1}(X) - |\gamma_P - 1|^L U_{L-2}(X) > 0$.
- (iii) If the lead time of the system, L , is odd, then the third condition simplifies to $\gamma_I > \max\{0, 2(\gamma_P - 1)\}$, and if the lead time of the system, L , is even, then the third condition simplifies to $0 < \gamma_I < 2$.

This conjecture has been verified numerically for $L = 2, \dots, 200$. Furthermore, the stability region appears to converge asymptotically. Given this, we can define a region of stability independent of the lead time. Formally, this asymptotic stability condition is defined by,

Lemma 1. *For all values of lead times $L \in \mathbb{N}$, stability is guaranteed within the region that is bounded by the lines $\gamma_I = 0$, $\gamma_I = 2$, $\gamma_I = 2(\gamma_P - 1)$, and $\gamma_I = 2\gamma_P$.*

This region is equivalent to that derived in Wei *et al.* ((2013), Proposition 4.3) through the analysis of the difference equations of the system. Moreover, it is independent of the parity of the lead time; i.e., whether it is odd or even. However, note that the difference we observe between odd and even lead times is a consequence of the mathematical analysis behind the results. To explain how this difference emerges in our model, we service [Conjecture 1](#). Unlike conditions (i) and (ii), only condition (iii) depends on the parity of the lead time. In particular, as per the proof of [Theorem 1](#), condition (iii) stems from the inequality $(-1)^{L+1} \Delta_{L+1} > 0$, which involves the evaluation of roots of a polynomial of order $L + 1$, where the parity of L affects the quality of the roots. For example, when L is even, we know that there is at least one real root, whereas for odd L the roots could be solely complex conjugates. Thus, [Lemma 1](#) confirms that there indeed should not be difference between even and odd lead times.

In the next section, we build upon the pole analysis used thus far and analyze the aperiodicity of the system.

3.2. Aperiodicity

If a system has a time-domain response with a number of maxima or minima that is less than n , the order of the system, we call such a system aperiodic (Jury, 1985). These dynamics are also defined by the poles of the transfer function: positive real poles contribute a damping component to the response, whereas negative real poles and poles with an imaginary component contribute oscillatory terms (Nise, 2011). Formally see the following definition.

Definition 3. (Jury, 1985). Suppose that $G(z) = N(z)/C(z)$ is the transfer function of a stable, linear, time-invariant system. Thus, all poles of the transfer function, p_i , $i = 1, \dots, n$, are within the unit circle. The response of this system is aperiodic if $\forall i, p_i \in [0, 1)$. From Disney (2008), we adopt the concept of a weakly aperiodic system if $\forall i, p_i \in \mathbb{R}$ and there exists an index $k \in \{1, \dots, n\}$ such that $p_k < 0$.

By analyzing the poles of the reduced characteristic polynomial (18) for APVIOBPCS systems and applying [Definition 3](#), we obtain the following propositions.

Proposition 2. *When $\gamma_I = \gamma_P = \gamma$, the response of a stable system for all lead times L is*

- aperiodic when $0 < \gamma \leq 1$, and
- weakly aperiodic when $1 < \gamma < 2$.

Proposition 3. *When $L > 2$ and $\gamma_I \neq \gamma_P$, the response of a stable system is non-aperiodic.*

We can define aperiodicity and weak-aperiodicity for the cases of $\gamma_I \neq \gamma_P$ and $L = 1, 2$. The area shaded in yellow represents the region for which the system is aperiodic in [Figure 1\(a\)](#) and weakly aperiodic in [Figure 1\(b\)](#). The boundaries for these regions can be found by following the same analysis as in the proof of [Proposition 3](#).

Remark 3. Pole analysis can be also used to determine whether the response of a given system is dampened or oscillatory. When the $L + 2$ poles of the transfer function (roots of the characteristic polynomial) reside within the unit circle and are real and positive, the response is dampened. When at least one of the poles is imaginary, or negative, the response is oscillatory. However, we cannot derive general statements on the performance of the system through pole analysis, due to the amount and magnitude of the poles being dependent on the order of the system and on the specific behavioral parameters.

The analysis of stability is a necessary condition for any study of an APVIOBPCS design, as a stable system guarantees bounded orders and inventories for any possible finite demand. Similarly, the pole analysis of the system is relevant because an aperiodic (or dampened) system avoids costly oscillations. By themselves, however, stability boundaries and pole analysis are not enough to measure the performance of the system under different demand conditions. The stability conditions and aperiodicity propositions, as well as the special regions defined in the accompanying figures, must be seen as a necessary first step in the evaluation of the system. In the next section, we perform extensive numerical experiments to evaluate performance metrics defined in [Section 2.1](#) as a function of the behavioral parameters of the system.

4. The impact of behavioral biases on performance

In this section, we study the impact of different behavioral biases on the stationary and non-stationary performance of an APVIOBPCS design. Closed-form expressions for the performance metrics introduced in [Section 2](#) can be derived for a limited range of parameter combinations (Hoberg *et al.*, 2007a; Hoberg and Thonemann, 2014). Due to the transcendental nature of the transfer functions of the system, however, it is not possible to do so when $\gamma_I \neq \gamma_P$, which is precisely the behavioral space that we are interested in. Hence, we perform numerical experiments to strengthen our understanding of the system. In terms of experimental setup, we fix the structural parameters of the system (with $\alpha = 0.3$, $C = 3$, $L = 5$) and vary γ_I and γ_P to quantify the performance over a range of different behavioral parameter settings (i.e., behavioral policies). We test the sensitivity of the system to changes in the structural parameters in [Appendix A](#). This analysis shows that in general, although

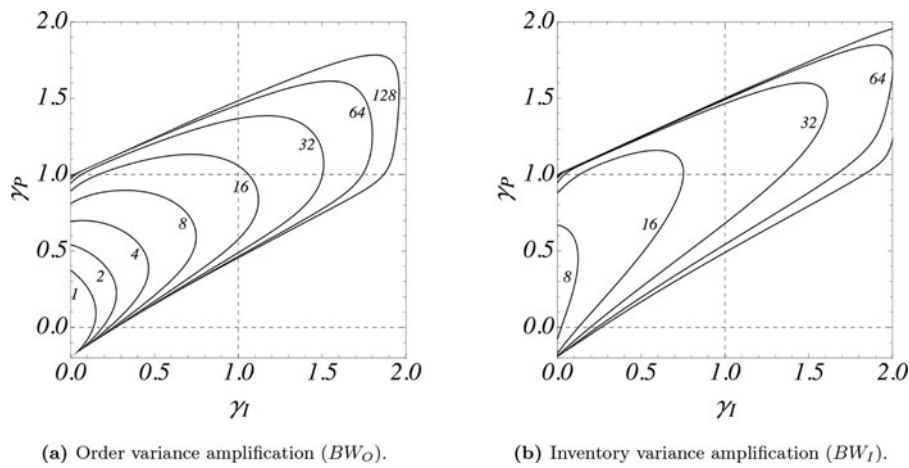


Figure 2. Stationary bullwhip contour plots.

the structure of the system affects its behavior, the insights discussed in the current section hold. In Section 4.1, we measure the stationary performance of the system in terms of two behavioral biases: smoothing/over-reaction to inventory and pipeline mismatches (Section 4.1.1 and Section 4.1.2) and under-/over-estimation of the pipeline (Section 4.1.3). Then, in Section 4.2, we excite the system with a one-time step change in the demand to measure its dynamic performance with respect to the aforementioned behavioral biases.

4.1. Stationary analysis

We use the metrics presented in Section 2.1.1 (amplification ratio), Section 2.1.2 (the bullwhip measure), and Section 2.1.3 (worst-case amplification) to understand how behavioral biases affect the performance of the system given stationary demand assumptions. The first behavior we study, over-reaction to inventory and pipeline mismatches, can be thought of as a bias born out of a panic reaction—a desire to reach the target level as soon as possible. The literature warns about the detrimental effects of such behavior under certain parameter settings; Disney *et al.* (2008) show that in the case of a DE-APVIOBPCS policy, $\gamma_I = \gamma_P > 1$ always induces a stationary bullwhip. In Section 4.1.1, we show that this holds for general APVIOBPCS systems; the bullwhip measure increases rapidly when $\gamma_I > 1$ or $\gamma_P > 1$. The converse behavior, under-reaction to mismatches (i.e., order smoothing), is a widely adopted strategy for bullwhip reduction (Disney *et al.*, 2008). Hence, in Section 4.1.2, we analyze the performance of the system within the “smoothing” behavioral region ($\gamma_I, \gamma_P < 1$). We find that smoothing reduces the bullwhip of both orders and inventories but that in terms of worst-case amplification it only has a significant impact on the orders. Worst-case inventory amplification appears relatively robust to smoothing. Furthermore, our analysis shows that the under-/over-estimation of the pipeline interacts with smoothing, so that it is not possible to ascribe, *a priori*, a positive or negative performance impact to such behavioral biases. Given that bullwhip contour lines span the entire behavioral region, any stationary-bullwhip target that we achieve with an unbiased (i.e., DE) policy we can also achieve with policies that under-estimate the pipeline as well as with policies that

over-estimate the pipeline. To better understand this behavioral bias, in Section 4.1.3, we analyze it through frequency response plots. In particular, we study the response to a series of experiments along and across contours to understand what drives the observed performance.

4.1.1. Over-reaction to inventory and pipeline mismatches

When $\gamma_I > 1$, the decision maker over-reacts to mismatches between the actual and desired inventory levels. For example, if $\gamma_I = 1.5$, then they will order 1.5 units for every 1 unit difference between \hat{i}_t and i_t in any given period t . Analogously, when $\gamma_P > 1$, the decision maker over-reacts to mismatches between the actual and desired pipeline levels (\hat{p}_t and p_t , respectively). To illustrate the effect of over-reaction, Figure 2 shows the values of BW_O and BW_I as a function of γ_I and γ_P . The bullwhip grows rapidly over the stable regions of the system when either parameter is larger than one. Hence, from a stationary perspective, there is no advantage in over-reacting to inventory and pipeline mismatches. In terms of relative performance, we see that order variance is particularly sensitive to over-reaction. A comparison between Figures 2(a) and 2(b) shows that the bullwhip of orders increases faster than the bullwhip of inventories when over-reaction is present.

4.1.2. Smoothing of inventory and pipeline mismatches

The under-reaction to inventory and pipeline mismatches is often referred to as order (or production) smoothing. In contrast with the over-reaction to mismatches, order smoothing is often a deliberate decision taken to reduce variability. Hoberg and Thonemann (2015), for example, showed that order smoothing diminishes the stationary bullwhip in DE-APVIOBPCS systems. The same is true for general APVIOBPCS policies. Figure 3 illustrates this by showing contour plots for the stationary bullwhip measures (BW_O and BW_I) for order-smoothing policies and expands on this by superimposing a density plot for the logarithm of the worst-case amplification metrics (A_O^∞ and A_I^∞). Figure 3(a) plots the metrics for orders as a function of the behavioral parameters and Figure 3(b) plots the equivalent for inventories. We see that low values of γ_I and γ_P correlate with low values of BW_O and BW_I and, as seen in Section 4.1.1,

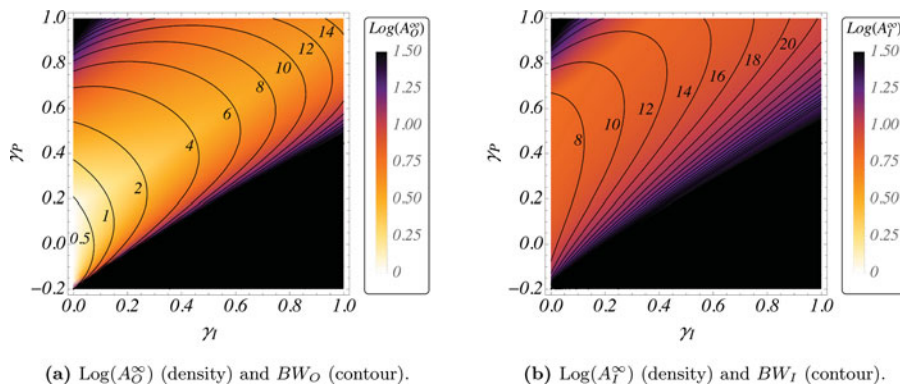


Figure 3. Worst-case amplification density plots and bullwhip contours. Black indicates $\log A^\infty \geq 2.5$.

increase from the lower left quadrant toward the upper right. Additionally, the contour plots show that any given value of BW_O or BW_I achieved by an unbiased policy can also be achieved by both over- and under-estimating biased policies. Thus, we cannot conclude that such biases by themselves are detrimental to the stationary performance of the system. Comparing Figures 3(a) and 3(b), we see that the BW_O response is more symmetric along the $\gamma_I = \gamma_P$ line than the response of BW_I . Moreover, BW_I is larger than BW_O for any given combination of behavioral parameters. In terms of the worst-case amplification, A_O^∞ is significantly more sensitive to behavioral biases than A_I^∞ . We observe that A_O^∞ increases markedly in γ_I and γ_P , but A_I^∞ does not show such an increase. Along any given contour, A_O^∞ is smallest at or close to the $\gamma_I = \gamma_P$ line, with a high dynamic range (i.e., the ratio between extreme values) along any given contour. For inventories, on the other hand, the minimum A_I^∞ appears above the $\gamma_I = \gamma_P$ line and its dynamic range along any given contour is smaller. As one would expect, worst-case amplification tends to infinity toward the stability boundaries (defined in Theorem 1) irrespective of the contour.

Under i.i.d. normally distributed demand, the analysis of BW_O and BW_I shows that heavy smoothing (low values of γ_I and γ_P) carries a performance advantage in terms of amplification; in fact, smoothing can generate values of $BW_O < 1$ under these settings. Under more general demand streams, however, the analysis of A_O^∞ and A_I^∞ shows that the orders of the system are more sensitive to behavioral biases than its inventory. This implies that the behavioral bias of under- and over-estimation of the pipeline has a particularly significant effect on the performance of the system when demand is not i.i.d. normally distributed.

4.1.3. Under-/over-estimation of the pipeline

To better understand the drivers behind the interaction between order smoothing and the under-/over-estimation of the pipeline, we analyze how the frequency response of the system varies along a given contour and how it differs across different contours within order-smoothing policies ($\gamma_I, \gamma_P < 1$). The bias of under-estimating the pipeline has long been identified as the main behavioral cause of the bullwhip effect (Sterman, 1989; Croson and Donohue, 2006; Croson et al., 2014). Over-estimating the pipeline has been linked to a comparable phenomenon (the reverse bullwhip effect) triggered by supply-side disruptions (Rong et al., 2008; Bueno-Solano and Cedillo-Campos, 2014).

Figure 4 shows the frequency response of orders under six different behavioral policies corresponding to two different BW_O contours: the three policies in Figure 4(b) correspond to $BW_O = 2$ and the three policies in Figure 4(c) correspond to $BW_O = 6$. Similarly, Figure 5 shows the frequency response of inventories under six different behavioral policies, grouped according to BW_I : the three policies in Figure 5(b) correspond to $BW_I = 12$ and the three policies in Figure 5(c) correspond to $BW_I = 16$. In each plot, we see a policy with over-estimation of the pipeline ($\gamma_I < \gamma_P$, dashed line), a policy with under-estimation of the pipeline ($\gamma_I < \gamma_P$, dashed-dotted line), and an unbiased policy ($\gamma_P = \gamma_I$, solid line).

As we observe from Figures 4 and 5, the influence of the pipeline bias on the amplification of orders and inventory variance is directly related to the harmonic signature of the demand; i.e., the contribution of single-frequency harmonics to the spectral density of the demand (Dejonckheere et al., 2003). In particular, the behavioral policies cause systems to

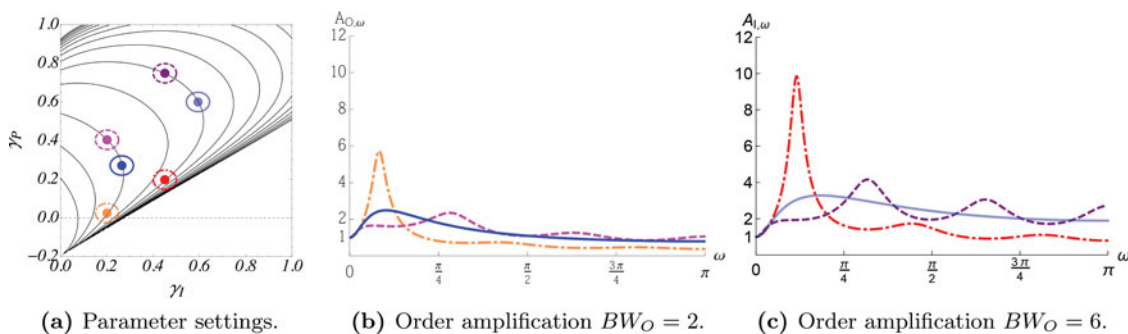


Figure 4. Bullwhip contour and frequency plots for orders.

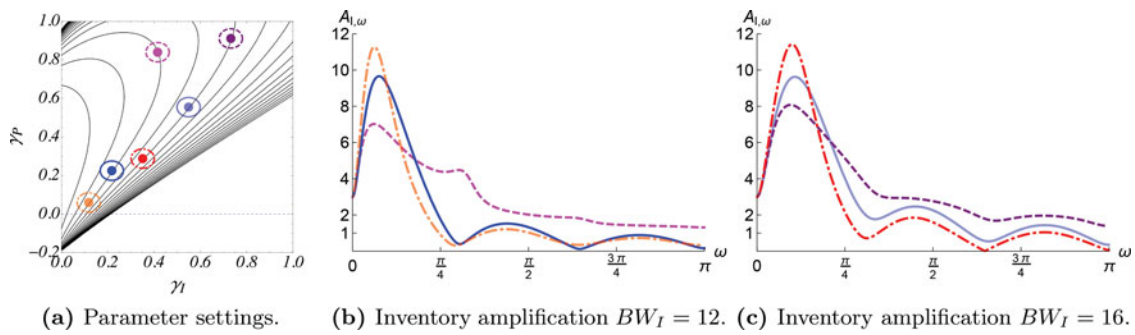


Figure 5. Inventory amplification contour and inventory frequency plots.

react differently to demands with predominantly high- and low-frequency harmonics, with a very clear tradeoff in performance. Under-estimating the pipeline (dashed-dotted lines) attenuates harmonics whose frequency is larger than roughly $\pi/4$ (if demand is observed daily, a frequency of $\pi/4$ represents a period of 8 days) but considerably amplifies harmonics of lower frequency. Conversely, over-estimating the pipeline (dashed lines) attenuates harmonics at frequencies lower than roughly $\pi/4$. A behavioral bias that attenuates a given harmonic frequency maximizes amplification at another frequency.

Unbiased policies (solid line) are, for any given harmonic frequency, outperformed by one of the biased policies (in some cases by both). However, the order response along the DE-diagonal is the most robust (i.e., flatter) of all. These observations complement the preceding analysis of the worst-case amplification metric. This metric only considers the maximum amplification peak across the entire frequency spectrum and can thus be directly computed from the frequency plots. Thus, the aforementioned robustness of the inventory worst-case amplification to behavioral biases follows immediately from an analysis of these plots. For both inventory and orders, the worst-case amplification occurs at relatively low harmonic frequencies, but in the case of inventories all pipeline estimation biases show similar peaks. This suggests that the inventory amplification is dominated by the harmonics in the demand, whereas the order amplification is sensitive to the combination of demand harmonics and behavioral bias.

It is important to note that the above analysis does not imply cyclical demands. Since any given demand pattern can be decomposed into a sum of single-frequency harmonics, these insights can be directly applied to any arbitrary demand pattern. Multiple methods exist to decompose any time series into its constituent harmonics; e.g., using Fast Fourier Transforms

(Cochran *et al.*, 1967) in MS Excel. Under-estimating (over-estimating) the pipeline will buffer demand patterns with high (low)-frequency peaks in their harmonic signature. In extreme cases of clearly cyclical demands, however, structural changes (i.e., seasonal forecasts) may be required to control amplification. In the next section, we investigate the performance of the system with respect to a special type of non-stationary demand: a one-time shock.

4.2. Dynamic analysis

To quantify the dynamic performance of the system, we use the ITAE metrics defined in Equations (14) and (15). Intuitively, we can relate the ITAE metric of a given system to its time-domain response; this metric is proportional to the area between the response curves and the steady state responses. Therefore, to build intuition behind the qualitative influence of behavioral biases on the dynamic response of the system, we present the time-domain response for a number of behavioral biases following a one-time demand shock in Figure 6. Figure 6(a) shows the evolution of orders and demand (dotted line) based on the experimental design of Figure 4(c). Figure 6(a) shows the evolution of inventory and target inventory (dotted line) based on the experimental design of Figure 5(c). This illustrates the qualitative influence of behavioral biases on the dynamic response of the system. Responses plotted with a dashed line correspond to over-estimation of the pipeline; those with a solid line correspond to a DE-response; and with a dashed-dotted line, to under-estimation of the pipeline. In general, over-estimating the pipeline causes a dampened response, whereas under-estimating the pipeline causes an oscillatory response (for any given policy, the response type can be predicted following Remark 3).

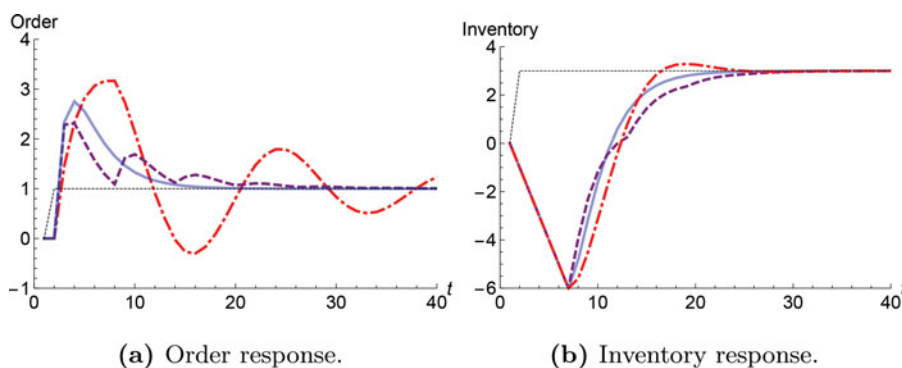


Figure 6. Time domain response to a step increase in demand.

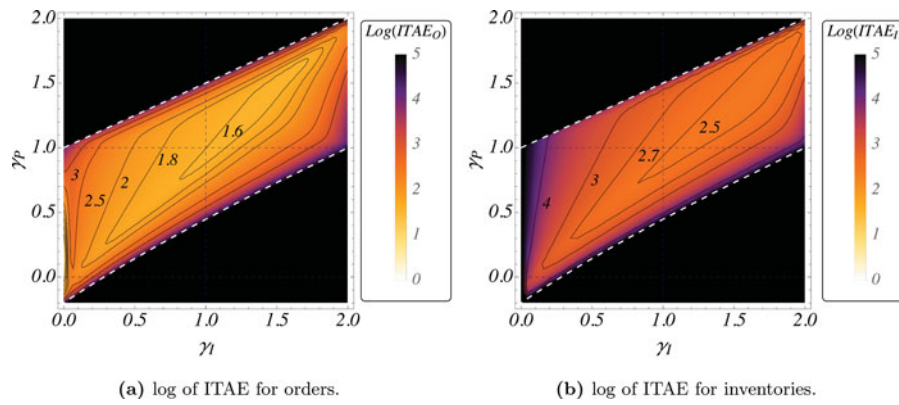


Figure 7. ITAE density plots. Black indicates $\log(\text{ITAE}) \geq 5$.

Figure 7 displays ITAE_O and ITAE_I as a function of γ_I and γ_P using a density plot. Numerically, we calculate the ITAE for 200 periods following the one-time shock and plot logarithmic values due to the extreme dynamic range. To better understand the influence of the behavioral parameters on the dynamic performance, we superimpose contours for a number of values. In contrast with the stationary metrics, we find that the dynamic performance of the system does not deteriorate rapidly with an over-reaction to orders nor improve with order smoothing. Rather, the dynamic performance of both orders and inventories appears to perform best around the classic base-stock policy area ($\gamma_I = \gamma_P = 1$). This observation is consistent with Hoberg and Thonemann (2015), who recommend such a policy when excellent dynamic performance is required. Outside this area, the performance is most robust close to the DE-diagonal ($\gamma_I = \gamma_P$) and deteriorates rapidly otherwise. Note, however, that the effect of smoothing and over-reacting is not symmetric along this diagonal. The dynamic performance deteriorates rapidly under heavy smoothing and heavy over-reaction, but the performance under moderate smoothing is more sensitive than under moderate over-reaction. This suggests that there is little incentive to adopt over-reaction in practice. Also note that the contours are closed lines. This indicates that if a given target ITAE can be achieved by one type of policy (say, a DE-policy), then it is possible to achieve the same performance with any other type of policy (smoothing, over-reacting to mismatches; under- and over-estimating the pipeline). Observing the time-domain response illustrated in Figure 6, we see that the decrease in performance due to under- or over-estimating the pipeline, although quantitatively similar, stems from contrasting behaviors. Over-estimating the pipeline degrades performance through over-dampened oscillations, whereas under-estimating the pipeline does so through under-dampened oscillations. Similarly, a given decrease in dynamic performance (ITAE) due to smoothing or over-reacting is driven by opposing behaviors. Under a smoothing (over-reacting) policy, orders and inventories converge gradually (rapidly) to the target values with minor (substantial) under- and over-shoots. Thus, for the decision maker, the dynamic tradeoff is between speed and variability.

At extreme values of γ , the performance of the system appears to break down. Dynamic performance in general decreases rapidly near the stability boundaries and inventory performance, in particular, decreases rapidly near the $\gamma_I = 0$ boundary. It can be confirmed, through pole analysis, that the amplitude and frequency of the oscillations increase toward

the stability boundaries defined by Conjecture 1. When $\gamma_I = 0$, the gap between actual and target inventories is not taken into account in the ordering equation, which causes the actual inventory to never approximate its target value. Hence, the poor inventory performance in such cases.

4.3. Stationary and transient performance tradeoff

From the previous sections, we know that the effect of behavioral biases on the performance depends on the underlying demand assumptions. Some policies (i.e., $\gamma_I = \gamma_P < 1$) increase the stationary performance but deteriorate the dynamic performance. Others (i.e., $\gamma_I = \gamma_P > 1$) have the opposite effect, and yet others (i.e., $\gamma_I < \gamma_P$) bring about non-trivial tradeoffs. In this section, we quantify the tradeoff between stationary and dynamic performance under different behavioral biases.

To visualize the tradeoff, we select three sets of stable behavioral policies (a DE-set, a supply line under-estimating set, and a supply line over-estimating set) and plot the logarithm of the stationary bullwhip metrics (BW_O and BW_I) against the logarithm of the dynamic metrics (ITAE_O and ITAE_I) for each set across the parameter space (i.e., from smoothing to over-reacting). We define the DE-set as $0.05 < \gamma_I = \gamma_P \leq 1.95$; the under-estimating set as $\gamma_I = \gamma_P + 0.15$ with $-0.1 \leq \gamma_P \leq 1.75$; and the over-estimating set as $\gamma_I = \gamma_P - 0.15$ with $0.2 \leq \gamma_P \leq 1.8$. Note that we exclude the extreme values of $\gamma_I = \{0, 2\}$ from the experimental design, due to their extremely poor dynamic performance. Figures 8(a) and 8(b) show the performance tradeoff curves of orders and inventories. As is the convention, the behavioral bias of under-estimating the pipeline is plotted with dashed-dotted lines; the bias of over-estimating the pipeline with dashed lines; and the unbiased pipeline (DE-policy) with a solid line. We direct the graphs through the use of different markers. The circle markers indicate the performance tradeoff for the setting closest to an OUT-policy for each set. ($\gamma_I = \gamma_P = 1$ for the DE-set; $\gamma_I = 1, \gamma_P = 0.85$ for the under-estimating set; and $\gamma_I = 1, \gamma_P = 1.5$ for the over-estimating set.) The triangle markers indicate the performance tradeoff of the setting with the largest stable smoothing for each set (i.e., $\gamma_I = 0.05$). The square markers indicate the performance tradeoff of the setting with the largest stable over-reaction for each set ($\gamma_P < 2$). Thus, the line segments between the triangle and circle markers indicate the performance of smoothing policies, and the line segments between circle and square markers indicate the performance of over-reacting policies.

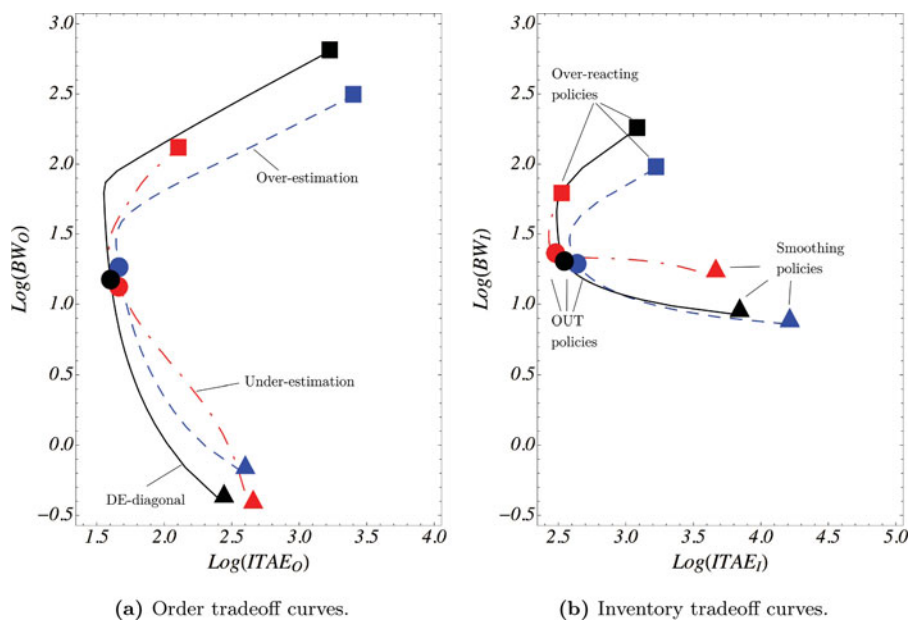


Figure 8. Performance tradeoff curves between stationary and transient responses.

The dominant policy from a tradeoff perspective is that which lies closest to the origin of the tradeoff plot. In this sense, smoothed DE-policies ($\gamma_P = \gamma_I < 1$) offer the best tradeoff between stationary and dynamic performance. However, in any given dimension, there exist behavioral biases that outperform the DE-policies. For example, the lowest $ITAE_I$ corresponds to policies that under-estimate the pipeline, and the lowest BW_I corresponds to a policy with over-estimation of the pipeline. Furthermore, the best dynamic performance for a given set is associated with over-reacting policies. This is consistent with the analysis of Section 4.2. However, the dynamic performance for orders is particularly sensitive to the over-reaction bias; performance deteriorates significantly once past the “sweet-spot” of over-reaction.

It is important to note that the plots presented here illustrate a tradeoff between two extreme forms of demand and, as such, only provide an intuition of the expected performance of a system confronted with a real demand stream. The plots, however, show that behavioral biases by themselves are not sufficient to make predictions about performance. Even an extreme over-reaction bias, which otherwise produces poor performance under every setting, results in relatively well-behaved dynamic inventory performance. Therefore, any analysis of human biases in the performance of an inventory system needs to explicitly consider the characteristics of the demand stream.

5. Conclusions

In this article, we used classic control theory to model general APVIOBPCS systems and analyze the impact of a number of behavioral biases (i.e., smoothing/over-reaction to target mismatches, and under-/over-estimation of the pipeline) on their stability and performance. Behavioral biases in this context are parametrized through feedback controllers γ_I and γ_P , which represent the incomplete closure of inventory and pipeline gaps at the moment of generating replenishment orders. Moreover, we showed how the behavioral biases affect the

stability of the system and developed a closed-form expression to determine the exact region of stability for any arbitrary lead time. This stability test has the advantage of avoiding the direct calculation of determinants or matrix-based procedures that characterize previous exact solutions of the problem (Jury, 1964; Disney, 2008). Additionally, this procedure allowed us to find an asymptotic region of stability independent of the lead time. This enables decision makers to operate in robust areas that guarantee stability independent of changes in the structural parameters. Through numerical experimentation, we showed that the performance of a stable system depends, to a large extent, on the behavioral biases but that this performance should not be analyzed independently of the demand stream.

We contribute to the bullwhip effect literature by explicitly modeling behavioral biases such as the under-estimation of the pipeline, long recognized as its main behavioral cause (Sternan, 1989). Although previous control-theoretic models have recognized the tradeoff between stationary and dynamic performance (Hoberg and Thonemann, 2014) as well as the use of independent controllers (Disney, 2008), our study is the first, to the best of our knowledge, to adopt this modeling methodology to explicitly link behavioral biases and demand attributes to system performance. Our dynamic performance results are largely consistent with prior behavioral research. We show that under-estimation of the pipeline degrades the system’s performance in the presence of demand shocks. We expand on this insight showing that the complementary bias, over-estimation of the pipeline, also has a negative effect under such conditions. However, we show that when the demand stream is stationary, the system is relatively robust to this bias. In such cases, we find biased policies (both under-estimating and over-estimating the pipeline) that perform just as well as unbiased policies (i.e., a DE-policy).

Order smoothing is prescribed as a strategy to limit the bullwhip effect (Disney, 2008). Empirical research, in fact, demonstrates that order smoothing and the bullwhip effect are concurrent in industry (Bray and Mendelson, 2015). We show that order smoothing is beneficial for the system’s performance when demand is stationary. Its impact, however, is

limited to the worst-case order amplification when demand is unpredictable. Dynamic analysis, on the other hand, reveals that order smoothing can degrade performance in the presence of demand shocks. The opposite bias (i.e., over-reaction to mismatches), on the contrary, degrades the stationary performance but can increase dynamic performance; controlled over-reaction can help the system achieve its new targets quickly. The system, however, is considerably sensitive to this behavior; excessive over-reaction significantly degrades performance. Given the above observations, we analyzed the influence of the behavioral biases in terms of the tradeoff between stationary and dynamic performance. We showed that unbiased policies offer generally good results under a large range of demand types. Such policies do not, however, result in the best performance under a particular criteria. We can always find a biased policy that outperforms an unbiased policy for any one performance metric.

Our findings have several implications for the study of the bullwhip effect. Although the bullwhip effect as a theoretical phenomenon makes no *a priori* assumptions on demand streams, in terms of measurement, an explicit recognition must be made between those causes that depend on demand assumptions, and those that do not. With respect to practical implications, our results suggest that there is no global optimum in terms of behavior. Hence, tradeoffs must be analyzed in terms of demand expectations and policy recommendations need to be made taking firm-specific performance priorities into account. In terms of theory, our results have potential methodological implications. Research on the field typically uses different demand assumptions based upon whether they are centered on structural or behavioral causes of the bullwhip effect. The former assume stationary demand streams (Chen and Lee, 2009), whereas the latter are based on experiments exploiting demand shocks (Croson *et al.*, 2014).

Considering the significant influence of the demand stream on the impact of behavioral biases, a direction for future research is to characterize normative policies, both structural and behavioral, that consider real-world demand scenarios. This would build on, and refine, the tradeoff analysis described in this article. Since demand observed in real life is neither purely stationary nor composed entirely of shocks, further research can consider the use of behavioral policies as a way to tailor the robustness of the system to realistic demand time series or even disruption scenarios. This could lead, for example, to recommendations that recognize the position of a system within a supply chain (i.e., the optimal behavioral policy for a retailer facing consumer demand would be different from that of its supplier, or its supplier's supplier) or recommendations that recognize different market segments (i.e., policy recommendations for process industries would be different from those for capital goods industries). Such a "real-life" approach would be consistent with, and complement, recent developments in the area that characterize system performance with additional dimensions. For example, in line with Hoberg and Thonemann (2015), additional research can use total cost performance as a complement to order and inventory variability.

Another area for further research concerns the (structural) forecasting assumptions of the system. In the presence of, for example, non-stationary or cyclical demands, the use of exponentially smoothed forecasts can be an additional source

of variability. Thus, a valuable research direction is to further analyze alternative forecasting methodologies in terms of the behavioral and the structural bullwhip effect. Li *et al.* (2014) provided the first steps in this direction through the control theoretic analysis of dampened trend forecasts within an APVIOBPCS-OUT design.

Acknowledgement

We are grateful to the three anonymous referees, the associate editor, and the department editor, Professor Jennifer Ryan, for their useful comments that helped us improve our article considerably.

Notes on contributors

Maximiliano Udenio is an Assistant Professor in the School of Industrial Engineering at Eindhoven University of Technology, The Netherlands. Dr. Udenio received his Ph.D. from the Eindhoven University of Technology in 2014. His research interests include inventory theory, network analysis, and sustainable supply chain management.

Eleni Vatamidou is a Post-Doctoral Researcher in the NEO group at Inria Sophia Antipolis, Méditerranée, France, and works on the MARMOTE (MARKovian MODELing Tools and Environments) project. She received her Ph.D. from the Department of Mathematics and Computer Science at Eindhoven University of Technology, The Netherlands, in June 2015. Her current research interests include the performance analysis of queueing, risk, and supply chain models.

Jan C. Fransoo is a Professor of Operations Management and Logistics in the School of Industrial Engineering at Eindhoven University of Technology in The Netherlands. His research addresses a variety of topics in supply chain management, with a current focus on container supply chains, emerging market logistics, and sustainability.

Nico Dellaert is an Associate Professor of Operations Management and Logistics in the School of Industrial Engineering at Eindhoven University of Technology in The Netherlands. His research addresses a variety of topics including vehicle routing, city logistics, health care planning, and capacity planning.

ORCID

Maximiliano Udenio  <http://orcid.org/0000-0002-2971-2795>

Jan C. Fransoo  <http://orcid.org/0000-0001-7220-0851>

Nico Dellaert  <http://orcid.org/0000-0003-2343-7574>

References

- Abramowitz, M. and Stegun, I. (1965) *Handbook of Mathematical Functions with Formulas Graphs and Mathematical Tables*, Dover, New York, NY.
- Bray, R.L. and Mendelson, H. (2015) Production smoothing and the bullwhip effect. *Manufacturing & Service Operations Management*, **17**(2), 208–220.
- Bueno-Solano, A. and Cedillo-Campos, M.G. (2014) Dynamic impact on global supply chains performance of disruptions propagation produced by terrorist acts. *Transportation Research Part E: Logistics and Transportation Review*, **61**, 1–12.
- Chen, F., Drezner, Z., Ryan, J.K. and Simchi-Levi, D. (2000) Quantifying the bullwhip effect in a simple supply chain: The impact of forecasting, lead times, and information. *Management Science*, **46**(3), 436–443.
- Chen, F., Ryan, J.K. and Simchi-Levi, D. (2000) The impact of exponential smoothing forecasts on the bullwhip effect. *Naval Research Logistics*, **47**(4), 269–286.
- Chen, L. and Lee, H.L. (2009) Information sharing and order variability control under a generalized demand model. *Management Science*, **55**(5), 781–797.

- Cochran, W.T., Cooley, J.W., Favon, D.L., Helms, H.D., Kaenel, R.A., Lang, W.W., Maling, G., Nelson, D.E., Rader, C.M. and Welch, P.D. (1967) What is the fast Fourier transform? *Proceedings of the IEEE*, **55**(10), 1664–1674.
- Crosron, R. and Donohue, K. (2006) Behavioral causes of the bullwhip effect and the observed value of inventory information. *Management Science*, **52**(3), 323–336.
- Crosron, R., Donohue, K., Katok, E. and Sterman, J. (2014) Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, **23**(2), 176–196.
- Dejonckheere, J., Disney, S., Lambrecht, M. and Towill, D. (2002) Transfer function analysis of forecasting induced bullwhip in supply chains. *International Journal of Production Economics*, **78**(2), 133–144.
- Dejonckheere, J., Disney, S., Lambrecht, M. and Towill, D. (2003) Measuring and avoiding the bullwhip effect: A control theoretic approach. *European Journal of Operational Research*, **147**(3), 567–590.
- Dejonckheere, J., Disney, S., Lambrecht, M. and Towill, D. (2004) The impact of information enrichment on the bullwhip effect in supply chains: A control engineering perspective. *European Journal of Operational Research*, **153**(3), 727–750.
- Deziel, D. and Eilon, S. (1967) A linear production-inventory control rule. *Production Engineer*, **46**(2), 93–104.
- Disney, S. (2008) Supply chain aperiodicity, bullwhip and stability analysis with Jury's inners. *IMA Journal of Management Mathematics*, **19**(2), 101–116.
- Disney, S., Farasyn, I. and Lambrecht, M. (2006) Taming the bullwhip effect whilst watching customer service in a single supply chain echelon. *European Journal of Operational Research*, **173**(1), 151–172.
- Disney, S. and Grubbström, R. (2004) Economic consequences of a production and inventory control policy. *International Journal of Production Research*, **42**(17), 3419–3431.
- Disney, S.M., Hoshiko, L., Polley, L. and Weigel, C. (2013) Removing bullwhip from Lexmark's toner operations. Paper presented at the Production and Operations Management Society Annual Conference, Denver, CO, 1–10.
- Disney, S., Lambrecht, M., Towill, D. and Van de Velde, W. (2008) The value of coordination in a two-echelon supply chain. *IIE Transactions*, **40**(3), 341–355.
- Disney, S. and Towill, D. (2002) A discrete transfer function model to determine the dynamic stability of a vendor managed inventory supply chain. *International Journal of Production Research*, **40**(1), 179–204.
- Disney, S. and Towill, D. (2003) On the bullwhip and inventory variance produced by an ordering policy. *Omega*, **31**(3), 157–167.
- Disney, S. and Towill, D. (2006) A methodology for benchmarking replenishment-induced bullwhip. *Supply Chain Management: An International Journal*, **11**(2), 160–168.
- Disney, S., Towill, D. and Van de Velde, W. (2004) Variance amplification and the golden ratio in production and inventory control. *International Journal of Production Economics*, **90**(3), 295–309.
- Disney, S., Towill, D. and Warburton, R. (2006) On the equivalence of control theoretic, differential, and difference equation approaches to modeling supply chains. *International Journal of Production Economics*, **101**(1), 194–208.
- Edghill, J. and Towill, D. (1990) Assessing manufacturing system performance: frequency response revisited. *Engineering Costs and Production Economics*, **19**(1), 319–326.
- Eppen, G.D. and Martin, R.K. (1988) Determining safety stock in the presence of stochastic lead time and demand. *Management Science*, **34**(11), 1380–1390.
- Forrester, J. (1958) Industrial dynamics: A major breakthrough for decision makers. *Harvard Business Review*, **36**(4), 37–66.
- Hoberg, K., Bradley, J. and Thonemann, U. (2007a) Analyzing the effect of inventory policies on the nonstationary performance with transfer functions. *IIE Transactions*, **39**(9), 911–924.
- Hoberg, K., Bradley, J. and Thonemann, U. (2007b) Analyzing the effect of the inventory policy on order and inventory variability with linear control theory. *European Journal of Operational Research*, **176**(3), 1620–1642.
- Hoberg, K. and Thonemann, U. (2014) Modeling and analyzing information delays in supply chains using transfer functions. *International Journal of Production Economics*, **156**, 132–145.
- Hoberg, K. and Thonemann, U. (2015) Analyzing variability, cost, and responsiveness of base-stock inventory policies with linear control theory. *IIE Transactions*, **47**(8), 865–879.
- Hosoda, T. and Disney, S. (2006) On variance amplification in a three-echelon supply chain with minimum mean square error forecasting. *Omega*, **34**(4), 344–358.
- Jakšič, M. and Rusjan, B. (2008) The effect of replenishment policies on the bullwhip effect: A transfer function approach. *European Journal of Operational Research*, **184**(3), 946–961.
- John, S., Naim, M. and Towill, D. (1994) Dynamic analysis of a WIP compensated decision support system. *International Journal of Management Systems Design*, **14**(4), 283–297.
- Jury, E. (1964) *Theory and Application of the z-Transform Method*, Volume 3. Wiley, New York, NY.
- Jury, E. (1985) A note on aperiodicity condition for linear discrete systems. *IEEE Transactions on Automatic Control*, **30**(11), 1100–1101.
- Lee, H., Padmanabhan, V. and Whang, S. (1997a) The bullwhip effect in supply chains. *Sloan Management Review*, **38**(3), 93–102.
- Lee, H., Padmanabhan, V. and Whang, S. (1997b) Information distortion in a supply chain: The bullwhip effect. *Management Science*, **43**(4), 546–558.
- Li, Q., Disney, S. and Gaalman, G. (2014) Avoiding the bullwhip effect using damped trend forecasting and the order-up-to replenishment policy. *International Journal of Production Economics*, **149**, 3–16.
- Marden, M. (1966) *The Geometry of Polynomials*, American Mathematical Society, Providence, RI.
- Marr, R. and Vineyard, G.H. (1988) Five -diagonal Toeplitz determinants and their relation to Chebyshev polynomials. *SIAM Journal on Matrix Analysis and Applications*, **9**(4), 579–586.
- Nise, N. (2011) *Control Systems Engineering*, Wiley, New York, NY.
- Ortega, M. and Lin, L. (2004) Control theory applications to the production-inventory problem: A review. *International Journal of Production Research*, **42**(11), 2303–2322.
- Ouyang, Y. and Daganzo, C. (2006) Characterization of the bullwhip effect in linear, time-invariant supply chains: Some formulae and tests. *Management Science*, **52**(10), 1544–1556.
- Rong, Y., Shen, Z.J.M. and Snyder, L.V. (2008) The impact of ordering behavior on order-quantity variability: A study of forward and reverse bullwhip effects. *Flexible Services and Manufacturing Journal*, **20**(1–2), 95–124.
- Sarimveis, H., Patrinos, P., Tarantilis, C.D. and Kiranoudis, C.T. (2008) Dynamic modeling and control of supply chain systems: A review. *Computers & Operations Research*, **35**(11), 3530–3561.
- Silvester, J. (2000) Determinants of block matrices. *The Mathematical Gazette*, **84**(501), 460–467.
- Simon, H. (1952) On the application of servomechanism theory in the study of production control. *Econometrica*, **20**(2), 247–268.
- Sterman, J. (1989) Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment. *Management Science*, **35**(3), 321–339.
- Struik, D. (1969) *A Source Book in Mathematics, 1200-1800*, Harvard University Press, Cambridge, MA.
- Su, X. (2008) Bounded rationality in newsvendor models. *Manufacturing & Service Operations Management*, **10**(4), 566–589.
- Towill, D. (1982) Dynamic analysis of an inventory and order based production control system. *The International Journal of Production Research*, **20**(6), 671–687.
- Tversky, A. and Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, **185**(4157), 1124–1131.
- Udenio, M., Fransoo, J. and Peels, R. (2015) Destocking, the bullwhip effect, and the credit crisis: Empirical modeling of supply chain dynamics. *International Journal of Production Economics*, **160**, 34–46.
- Vassian, H. (1955) Application of discrete variable servo theory to inventory control. *Journal of the Operations Research Society of America*, **3**(3), 272–282.
- Warburton, R., Disney, S., Towill, D. and Hodgson, J. (2004) Technical Note: Further insights into “the stability of supply chains.” *International Journal of Production Research*, **42**(3):639–648.

- Wei, Y., Wang, H. and Qi, C. (2013) On the stability and bullwhip effect of a production and inventory control system. *International Journal of Production Research*, **51**(1), 154–171.
- Wright, D. and Yuan, X. (2008) Mitigating the bullwhip effect by ordering policies and forecasting methods. *International Journal of Production Economics*, **113**(2), 587–597.
- Wunsch, A. (2005) *Complex Variables with Applications*, Pearson, Boston, MA.
- Zhang, X. (2004) The impact of forecasting methods on the bullwhip effect. *International Journal of Production Economics*, **88**(1), 15–27.
- Zhou, L., Disney, S. and Towill, D. (2010) A pragmatic approach to the design of bullwhip controllers. *International Journal of Production Economics*, **128**(2), 556–568.

Appendices

Appendix A: Sensitivity to structural parameters

In this section, we test the sensitivity of the system to changes in the structural parameters α , C , and L . Figure A1 shows the sensitivity of the frequency response and Fig. A2 shows the sensitivity of the bullwhip measure. Each sub-figure corresponds to a value of C and contains nine plots arranged in a 3×3 matrix, each row containing plots with fixed α and $L = 1, 5, 10$ and each column containing plots with fixed L and $\alpha = 0.1, 0.3, 0.5$. Sub-figures on the left-hand side of the page represent order-related metrics and sub-figures on the right-hand side of the page represent inventory-related metrics.

We use the experimental design from Section 4.1.3 to generate the frequency response plots in Fig. A1. Figures A1(a), A1(c), and A1(e) are based upon the parametrization shown in Fig. 4(a), with $BW_O = 6$. Figures A1(b), A1(d), and A1(f) are based upon Fig. 5(a), with $BW_I = 16$.

We see that the inventory coverage, C , has a very limited impact on the amplification of the system. Increasing its value sixfold has a marginal effect on the absolute value of $A_{O,\omega}$ and $A_{I,\omega}$ and no effect on the shape of the response curves. The effect of the smoothing parameter, α , is more pronounced in terms of the absolute value of $A_{O,\omega}$ and $A_{I,\omega}$; we observe an increase in amplification at all frequencies as a response to an increase in α . However, a change in α also has no discernible qualitative effect on the frequency response. Lead time L , however, has both a qualitative and quantitative effect on the frequency response. The absolute value of $A_{O,\omega}$ and $A_{I,\omega}$ increases with an increase in L . Moreover, the shape of the response curves changes as L increases. In particular, the number and location of the high-frequency response peaks seem to be determined by the lead time. This effect is clearly seen in the frequency response plots of orders. An analysis of Fig. A1, however, suggests that the insights related to the behavioral biases do not depend on the structural parameters. The relative responses of under-/over-estimating policies are maintained throughout.

In Fig. A2 we extend the experiments shown in Section 4.1.1. Figures A2(a), A2(c), and A2(e) show the effect of the structural parameters on $\log(BW_O)$. Correspondingly, Figs. A2(b), A2(d), and A2(f) show the effect of the structural parameters on $\log(BW_I)$. The logarithmic scale used is the same as that shown in Fig. 2.

Analysis of Fig. A2 suggests that all of the structural parameters have a quantitative effect in the bullwhip measure. Similar to the analysis of the frequency response, we see that the effect of

C is comparatively minor to the effect of α and L . Finally, these additional experiments suggest that the insights on the behavioral biases developed in this article do not depend on the structural parameters.

Appendix B: Proofs

Proof of Proposition 1. We denote each root of the characteristic polynomial (pole of the transfer function) by p_i^L , where $i = \{1, 2, \dots, L + 2\}$.

It follows from Equation (17) that $p_1^L = (1 - \alpha)$ is a real root of the characteristic polynomial that does not depend on L .

When $\alpha \in (0, 2)$, p_1^L is inside the unit circle and the remaining $L + 1$ roots of the characteristic polynomial $C(z)$ will be equal to the $L + 1$ roots of the reduced characteristic polynomial $\hat{C}(z)$. Thus, the condition for stability when $\alpha \in (0, 2)$ reduces to checking that all roots of $\hat{C}(z)$ be inside the unit circle.

When $\alpha = 0$ or $\alpha = 2$, then $|p_1^L| = 1$, which means that the system will be marginally stable unless $|p_1^L| = 1$ is also a root of the numerator of the transfer function. The transfer function for orders when $\alpha = 0$ can be rewritten as

$$G_O(z) = \frac{\gamma_I(z-1)z^L}{(z-1)C\hat{C}(z)} = \frac{\gamma_I z^L}{C\hat{C}(z)}. \quad (\text{A1})$$

The transfer function for inventories when $\alpha = 0$ can be rewritten as

$$\begin{aligned} G_I(z) &= \frac{\gamma_I - z(z^{L+1} + z^L(\gamma_P - 1) - \gamma_P)}{(z-1)C\hat{C}(z)} \\ &= \frac{\gamma_I - z(z^L + \gamma_P z^{L-1} + \dots + \gamma_P z + \gamma_P)}{C\hat{C}(z)}. \end{aligned} \quad (\text{A2})$$

Thus, since $z = 1$ is a root of the denominator of both $G_O(z)$ and $G_I(z)$, $\forall L \in \mathbb{N}$, the conditions for stability when $\alpha = 0$ reduce to checking that all roots of $\hat{C}(z)$ be inside the unit circle. The same reasoning can be extended for the case of $\alpha = 2$. \square

Proof of Proposition 2. When $\gamma_I = \gamma_P = \gamma$ we can rewrite the reduced characteristic polynomial:

$$\hat{C}(z) = z^L(z - 1 + \gamma). \quad (\text{A3})$$

Its $L + 1$ zeroes are

$$p_2^L = (1 - \gamma), \quad (\text{A4})$$

$$p_3^L = p_4^L = \dots = p_{L+2}^L = 0. \quad (\text{A5})$$

When $\gamma \in (0, 2)$, $|p_2^L| < 1$; therefore, the system is stable. More precisely, for $\gamma \in (0, 1]$, $p_2^L \geq 0$, and for $\gamma \in (1, 2)$, $p_2^L < 0$. Thus, the system is respectively aperiodic and weakly aperiodic. \square

Proof of Proposition 3. We know that all of the roots of the reduced characteristic polynomial $\hat{C}(z)$ of a stable system lie inside the unit circle of the complex plane. To judge the aperiodicity of such a system, we need to know whether the roots of $\hat{C}(z)$ lie on the negative or positive half plane.

To do so, we rewrite the characteristic polynomial:

$$\hat{C}(z) = z^{L+1} - z^L(1 - \gamma_P) + \gamma_I - \gamma_P, \quad (\text{A6})$$

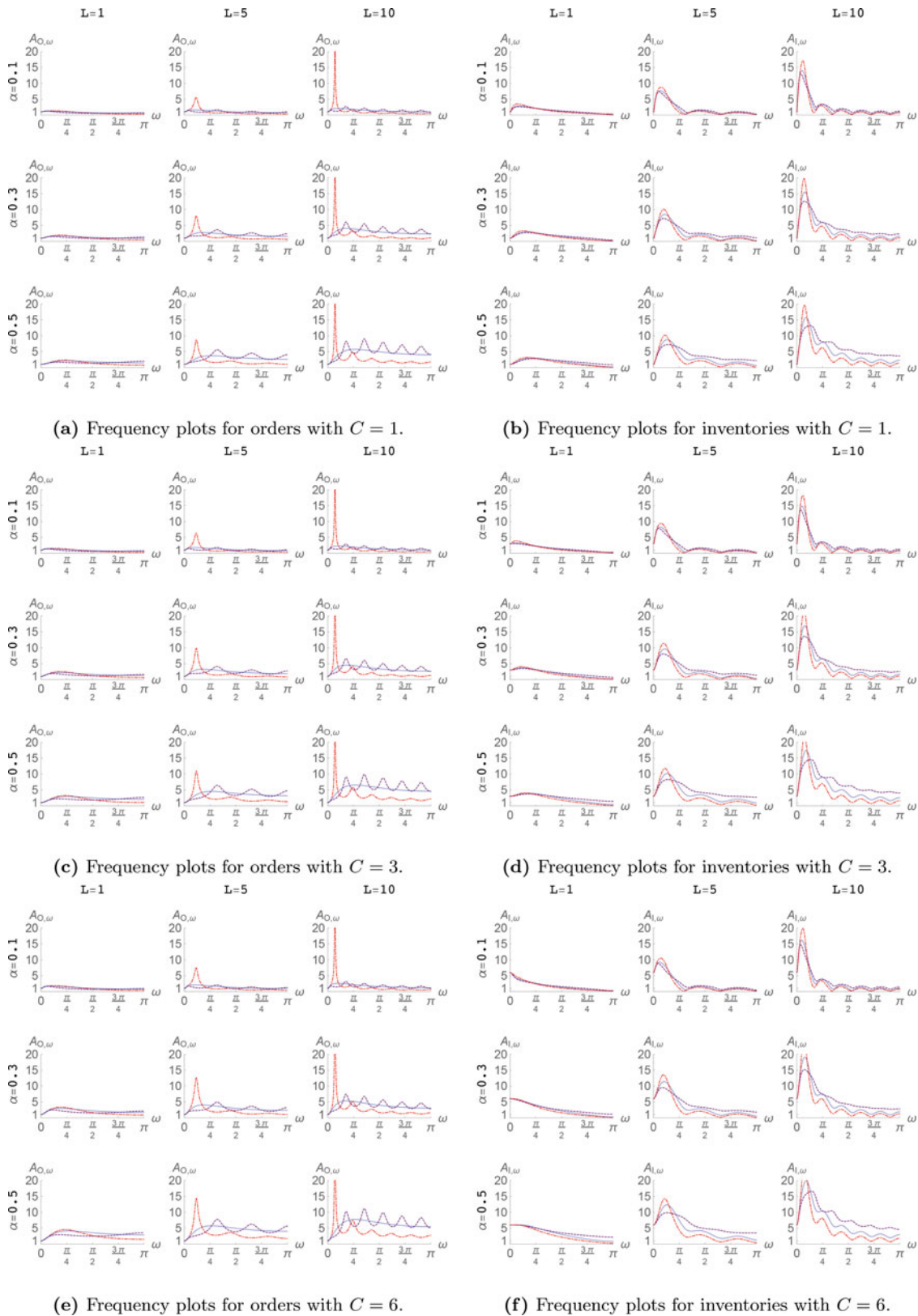


Figure A1. Sensitivity of frequency plots of orders (left) and inventories (right) to structural parameters.

and apply Descartes' rule of signs (Struik, 1969, pp. 90–94) to identify the number of positive and negative roots. We assume that the system is stable and distinguish between two cases: $\gamma_I < \gamma_P$ and $\gamma_I > \gamma_P$.

The case of $\gamma_I < \gamma_P$: For all values of L and γ_P this polynomial will have one sign change. Therefore, we will always have one

positive and real root. To find the negative and real roots of $\hat{C}(z)$, we separate between odd and even lead times L : For L odd, the polynomial $\hat{C}(-z) = z^{L+1} + z^L(1 - \gamma_P) + \gamma_I - \gamma_P$ has one sign change for all values of γ_P . Therefore, there exists a real and negative root of $\hat{C}(z)$ and the remaining $L - 1$ roots come in pairs of complex conjugates. Only for $L = 1$ does $\hat{C}(z)$ have no

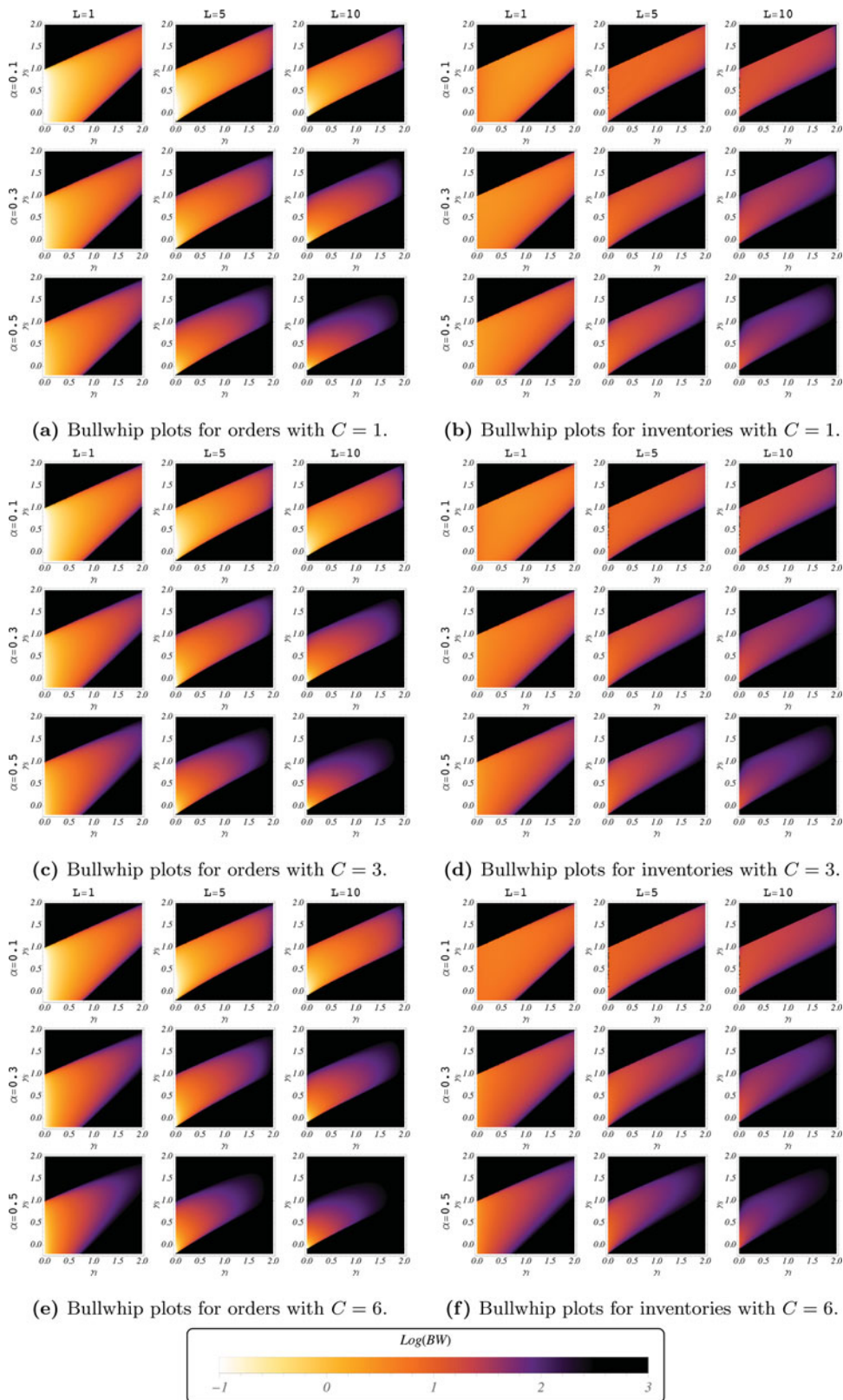


Figure A2. Sensitivity of the bullwhip for orders (left) and inventories (right). Log scale.

complex roots. For L even, the polynomial $\hat{C}(-z) = -z^{L+1} - z^L(1 - \gamma_P) + \gamma_I - \gamma_P$ does not have any sign change when $\gamma_P \in [0, 1]$ and thus no negative and real roots. This means that it has at least one pair of conjugate complex roots. When

$\gamma_P \in (1, 2)$, it has two sign changes and thus zero or two negative and real roots. If it has two negative roots and $L = 2$, then the system is weakly aperiodic. In any other case, there exists at least one pair of complex roots and the system is thus aperiodic.

The case of $\gamma_I > \gamma_P$: For all values of L , and for $\gamma_P \in [1, 2)$, the reduced characteristic polynomial $\hat{C}(z)$ will have no sign changes and consequently it does not have any positive and real roots. To find the negative and real roots of $\hat{C}(z)$ with $\gamma_P \in [1, 2)$, we separate between odd and even lead times L : For L odd, the polynomial $\hat{C}(-z)$ has two sign changes and therefore either two or zero real and negative roots. When $L = 1$ and it has two negative real roots, the system is weakly aperiodic. In all other cases, there exist at least one pair of complex roots and the system will consequently be non-aperiodic. For L even, the polynomial $\hat{C}(-z)$ has one sign change and therefore one real and negative root. Thus, in this case, the polynomial will always have at least one pair of complex roots and the system will consequently be non-aperiodic. For all values of L , and for $\gamma_P \in [0, 1)$, the reduced characteristic polynomial $\hat{C}(z)$ will have two sign changes and consequently it has either two or zero positive and real roots. To find the negative and real roots of $\hat{C}(z)$ with $\gamma_P \in [0, 1)$, we separate once more between odd and even lead times L : For L odd, the polynomial $\hat{C}(-z)$ has zero sign changes and therefore either zero real and negative roots. When $L = 1$ and it has two positive real roots, the system is aperiodic. In all other cases there exist at least one pair of complex roots and the system will consequently be non-aperiodic. For L even, the polynomial $\hat{C}(-z)$ has one sign change and therefore one real and negative root. Only the combination $L = 2$ and two positive real roots gives a weakly aperiodic response. In all other cases, there exists at least two pair of complex roots and thus the system is non-aperiodic. \square

Proof of Theorem 1. According to Theorem 43.1 of Marden (1966, p. 198), the number of roots of our reduced characteristic polynomial $\hat{C}(z)$ (Equation 18) inside the unit circle is equal to the number of negative signs in the sequence

$$\Delta_1, \frac{\Delta_2}{\Delta_1}, \dots, \frac{\Delta_{L+1}}{\Delta_L}, \quad (\text{A7})$$

where

$$\Delta_n := \det \begin{bmatrix} A_n & A_n^{*T} \\ A_n^* & A_n^T \end{bmatrix}, \quad (\text{A8})$$

and

$$A_n := \begin{bmatrix} a & 0 & 0 & \dots & 0 \\ 0 & a & 0 & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & a \end{bmatrix}, \quad n = 1, \dots, L,$$

$$A_{L+1} := \begin{bmatrix} a & 0 & 0 & \dots & 0 \\ 0 & a & 0 & \dots & 0 \\ 0 & 0 & a & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ b & 0 & 0 & \dots & a \end{bmatrix},$$

$$A_n^* := \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ b & 1 & 0 & \dots & 0 \\ 0 & b & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}, \quad n = 1, \dots, L+1,$$

with $a = \gamma_I - \gamma_P$, and $b = \gamma_P - 1$. Here A_n^T denotes the transpose of A_n where the dimension of these matrices is $n \times n$.

To guarantee stability, we need all of the roots of Equation (18) to be inside the unit circle. Thus, we need to have $L+1$ negative signs in the sequence (A7). Consequently, we need to have

$$(-1)^n \Delta_n > 0, \quad \forall n = 1, \dots, L+1. \quad (\text{A9})$$

Since the matrices A_n and A_n^* commute, according to Silvester (2000), we have that for $n = 1, \dots, L+1$, $(-1)^n \Delta_n = \det(A_n A_n^T - A_n^* A_n^{*T})$. Thus, for $n = 1, \dots, L$,

$$(-1)^n \Delta_n = \det \begin{bmatrix} 1 - a^2 & b & 0 & \dots & \dots & -ab \\ b & 1 - a^2 + b^2 & b & 0 & \dots & 0 \\ 0 & b & 1 - a^2 + b^2 & b & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & b & 1 - a^2 + b^2 & b \\ -ab & 0 & \dots & 0 & b & 1 - a^2 + b^2 \end{bmatrix}, \quad (\text{A10})$$

and also

$$(-1)^{L+1} \Delta_{L+1} = \det \begin{bmatrix} 1 - a^2 & b & 0 & \dots & \dots & -ab \\ b & 1 - a^2 + b^2 & b & 0 & \dots & 0 \\ 0 & b & 1 - a^2 + b^2 & b & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & b & 1 - a^2 + b^2 & b \\ -ab & 0 & \dots & 0 & b & 1 - a^2 \end{bmatrix}. \quad (\text{A11})$$

If we denote M_n as the square $n \times n$ matrix with diagonal elements equal to $1 - a^2 + b^2$ and all elements on the upper and lower diagonal are equal to b , then we can find recursively that

$$(-1)^n \Delta_n = (1 - a^2) \det M_{n-1} - b^2 \det M_{n-2}, \quad n = 2, \dots, L. \tag{A12}$$

The determinant $\det M_n$ can be calculated through formula (3) of Marr and Vineyard (1988) as

$$\det M_n = D_n(1 - a^2 + b^2, b, b) = |b| U_n \left(\frac{1 - a^2 + b^2}{2|b|} \right), \tag{A13}$$

where U_n is the n th degree Chebyshev polynomial of the second kind, defined by

$$U_n(Z) = \frac{(Z + \sqrt{Z^2 - 1})^{n+1} - (Z - \sqrt{Z^2 - 1})^{n+1}}{2\sqrt{Z^2 - 1}}. \tag{A14}$$

If we set $X = (1 - a^2 + b^2)/(2|b|)$, then

$$(-1)^n \Delta_n = (1 - a^2) |b|^{(n-1)} U_{n-1}(X) - |b|^n U_{n-2}(X). \tag{A15}$$

Similarly, for the $(L + 1)$ st determinant, it holds that

$$\begin{aligned} (-1)^{L+1} \Delta_{L+1} &= (1 - a^2)^2 \det M_{L-1} - 2(1 - a^2)b^2 \det M_{L-2} \\ &\quad + b^4 \det M_{L-3} + 2(-1)^{L+1} ab^{L+1} - (ab)^2 \det M_{L-1}, \end{aligned} \tag{A16}$$

which reduces to Equation (23). Observe that $-\Delta_1 = 1 - a^2$, which completes the proof. \square

Proof of Lemma 1. In a compact form, the region defined by Lemma 1 can be written as $|b| \leq 1 - |a|$ with $a = \gamma_I - \gamma_P$ and $b = \gamma_P - 1$. We observe that condition (iii) of Conjecture 1 always defines two boundary lines in this region. Therefore, inside this region, condition (iii) is always going to be satisfied. In order to show that this is indeed the asymptotic region defined by Lemma 1 when L goes to ∞ , it is sufficient to show that when we set $|b| = 1 - |a|$:

$$\lim_{L \rightarrow +\infty} (-1)^L \Delta_L = 0. \tag{A17}$$

Knowing that $X = 1$ (see proof of Theorem 1) and, using that $U_L(1) = L + 1$ (Abramowitz and Stegun, 1965, Table 22.3.7, p. 774), we can rewrite Equation (A15) as

$$(-1)^L \Delta_L = (1 - |a|)^L (1 + L|a|), \tag{A18}$$

which goes to zero as L goes to ∞ since $|a| < 1$, and the proof is complete. \square