Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The effects of quantity discounts on supply chain performance: Looking through the Bullwhip lens

Borja Ponte a, Julio Puche b,*, Rafael Rosillo a, David de la Fuente a

a Department of Business Administration, University of Oviedo, Gijón, Spain
b Department of Applied Economics, University of Burgos, Burgos, Spain

ARTICLE INFO

Keywords:
Bullwhip Effect
Inventory control
Purchasing
Quantity discounts
Supply chain

ABSTRACT

Quantity discounts are a common pricing mechanism to stimulate large orders. We explore their impact on the dynamic behaviour of production and distribution systems by studying key operational and economic metrics. In a three-echelon supply chain, we observe that the discount generally increases the Bullwhip Effect, which especially harms the manufacturer. The discount also reduces the retailer’s purchase costs, but increases its inventory- and capacity-related costs. A key trade-off thus emerges, which manifests itself through a U-shaped relationship between the total cost and the discount acceptance parameter. In the light of this trade-off, we discuss how key factors should affect the retailer’s willingness to pursue the discount. We observe that managers that need to deal with tougher environmental conditions, such as high demand uncertainty and long lead times, should be less reluctant to increase orders up to the discount quantity. We also discuss in detail other valuable insights for professionals, both from the perspective of sellers and buyers.

1. Introduction

Supply chains are victims of a dynamic phenomenon known as the Bullwhip Effect (Lee et al., 1997a, 1997b). This refers to the amplification of the variability of orders as they pass through the different echelons of a supply chain, which may also have negative implications in terms of the variability of inventories. On the whole, the Bullwhip Effect creates a climate of instability in production and distribution systems that significantly decreases their operational and financial performance (Metters, 1997; Disney and Lambrecht, 2008; Dominguez et al., 2018). Given its practical importance, the Bullwhip phenomenon has become a fruitful area of research over the last two decades; see e.g. the review of the Bullwhip literature by Wang and Disney (2016). However, the Bullwhip problem is still far from being solved. A few years ago, a study by Isaksson and Seifert (2016) reported a mean increase of the coefficient of variation (c.v.) of the orders that equalled 90% after analysing a sample of approx. 15,000 buyer–supplier dyad relationships in a wide range of US industries.

As is obvious by its definition, the Bullwhip Effect particularly affects the higher echelons of the supply chain, such as manufacturers. To prevent it from aggressively propagating upstream and thus damaging the operations of all supply chain partners, managers need to understand why it occurs. Disney and Lambrecht (2008) highlighted that this phenomenon emerges through the interaction of behavioural and operational components. The behavioural causes of the Bullwhip problem, studied in detail by Croson and Donohue (2006), derive from the fact that managerial decisions are not always completely rational; rather, decision makers commonly over- or
under-react to demand changes. In this fashion, managers repeatedly underweight the supply chain when issuing orders and their forecasts are often biased by overconfidence, as observed by Sterman (1989) and Ancarani et al. (2016) through Beer Game experiments. The operational causes of the Bullwhip Effect mainly refer to sources of information distortion throughout the supply chain. Lee et al. (1997a) identified four main operational causes that exist even when supply chain actors take perfectly rational decisions: the processing of the demand signal\(^1\), order batching\(^2\), rationing and shortage gaming\(^3\), and price fluctuations.

In relation to the last operational cause, Lee et al. (1997a) highlighted that some pricing strategies that have a positive marketing effect on companies, such as promotions, might significantly deteriorate the operational performance of supply chains. These incentives to buy more than what is actually needed at a specific point in time result in the fact that the customers’ buying pattern does not reflect their requirements. This significantly adds to the distortion of the relevant information throughout the supply chain, and hence contributes to losing control of the variability amplification. Due to their negative effects on the operational behaviour of organisations, such pricing strategies may become ‘the dumbest marketing ploy’, as described by Sellers (1992). We refer interested readers to Butman (2002) for an interesting discussion on common issues in the interface between marketing and operations management in the practice of organisations.

Along these lines, pricing decisions are well known to have a major impact on the operations of organisations and the dynamics of supply chains. Indeed, they may unintentionally contribute to the generation of costly inefficiencies, such as those derived from the Bullwhip Effect. However, price considerations have been rarely contemplated in the Bullwhip literature so far, probably due to the complexity of bringing them into the mathematical study of operations, as discussed by Bhattacharya and Bandyopadhyay (2011) and Wang and Disney (2016). The main findings of the relatively scarce literature that examines the effects of pricing on the Bullwhip phenomenon are considered below.

1. **Price repercussions on the Bullwhip Effect: A brief review of the literature**

As discussed earlier, Lee et al. (1997a) listed price fluctuations as one of the four operational causes of the Bullwhip Effect. This was demonstrated in a subsequent paper, Lee et al. (1997b), through a simple mathematical model. In the light of this finding, the authors proposed the use of everyday low pricing (EDLP) strategies, without running sales or price promotions, as a solution for the demand amplification problem. Almost a decade later, Gavirneni (2006) studied the same problem when information transparency exists in the supply chain. They observed that under such circumstances price fluctuations, e.g. in high-low (HL) pricing strategies, may lead to reduced supply chain costs. Similar findings were reported by Hamister and Suresh (2008). Specifically, they revealed that, under auto-regressive demand, the use of constant pricing may also result in lower profitability. Later, Zhang and Burke (2011) showed that marketing strategies based on price fluctuations may exacerbate or alleviate the Bullwhip Effect, depending on the auto- and cross-correlation between demand and price. Sodhi et al. (2014) studied the problem in a maintenance-repair-and-overhaul (MRO) setting. By incorporating discrete stochastic prices in an economic order quantity (EOQ) model, they concluded that the Bullwhip Effect grows as price variance increases. Finally, it is interesting to note that Su and Geunes (2012) discussed that, although price variations can increase operational costs, this may be more than offset by increased revenues, regardless of the existence of coordination mechanisms.

From these prior studies, we conclude that the relationship between the operational performance of supply chains and the variability of prices is not straightforward. Rather, this relationship significantly depends on the context under study, including relevant factors such as the presence or lack of collaborative mechanisms in the supply chain (e.g. information sharing), the serial correlation in the demand and price time series, and the cross-correlation between them. In this fashion, price variability sometimes provokes a sharp decrease in the operational performance of supply chain actors by means of a pricing-induced Bullwhip phenomenon (e.g. Lee et al. 1997b; Sodhi et al., 2014). However, on other occasions, supply chains may benefit significantly from strategically inducing changes in the prices of products. This occurs because, under specific circumstances, price variability acts as a Bullwhip-dampening mechanism (e.g. Hamister and Suresh, 2008; Zhang and Burke, 2011), while, in other cases, the performance of the supply chain improves even though the Bullwhip Effect grows. In the words of Su and Geunes (2012, p. 892), ‘the Bullwhip Effect might [sometimes] be a necessary evil in a profit-maximizing supply chain’. From this perspective, the authors suggested that firms would not focus on mitigating Bullwhip when it jeopardises other strategic objectives such as market share or total revenue.

Not many other studies (if any), apart from those just discussed, have investigated the Bullwhip Effect with price considerations. Indeed, in their contemporary review of Bullwhip studies, Wang and Disney (2016) identified six primary research gaps in this field that should be soon addressed. One of them is the study of ‘the influence of prices on Bullwhip’ (Wang and Disney, 2016, p. 697). It should be highlighted that, as evidenced by their review and our analysis in this section, these previous works have focused on

---

1. Often labelled as the Forrester Effect (see e.g. Disney and Towill, 2003), given that Jay W. Forrester (1958) identified four decades before the amplification effect of orders in supply chains due to the order-making processes.
2. Also known as the Burbidge Effect (see e.g. McCullen and Towill, 2002). Burbidge (1984) observed the Bullwhip Effect in UK shop floors, and attributed it to the discontinuous and ‘sharp edged’ nature of production, often governed by batching processes.
3. Sometimes named as the Houlihan Effect (see e.g. Kristianto et al., 2012). Houlihan (1985) highlighted that, when shortages occur in supply chains, customers tend to over-load their orders, which significantly contributes to amplifying the variability of orders.
4. Similar arguments were made by other authors. In a previous Bullwhip review, Bhattacharya and Bandyopadhyay (2011, p. 1254) noted that the ‘effect of pricing has more economic implications than those […] I considered in the existing literature’. Also, Cantor and Katok (2012) pointed out to the pricing effects on Bullwhip as a key area for future research.
understanding the effects of price variations by contrasting fixed-price strategies, such as EDLP, versus fluctuating-price strategies, such as HL. As we will discuss in the following subsection, the present paper differs from these previous works in the Bullwhip literature in that we study the relationship between pricing and Bullwhip by assessing the impact of quantity discounts.

1.2. Linking quantity discounts and performance through the Bullwhip effect: Our contribution

A quantity discount is an economic incentive to encourage organisations (or individuals) to purchase goods in large quantities. Basically, the vendor rewards buyers who make big purchases by providing a reduced unit price. These discounts are very common in practice, taking on a variety of formulas; see Weng and Wong (1993) and Rubin and Benton (2003). Maybe the most common formula is the all-unit discount, based on applying the lower unit price to all the units purchased when a certain threshold is reached (e.g., Chen and Robinson, 2012; Zhang et al., 2019). For instance, assume the regular price of a product is 1€. The seller may opt to offer a 10% discount that goes into effect if at least 100 units are purchased. That is, buyers who purchase less than 100 units would pay 1€ per unit, while buyers purchasing larger quantities would pay €0.9 per unit.

Quantity discounts not only may help the seller increase the sales revenue and reach economies of scale (Viswanathan and Wang, 2003), but also have other advantages for businesses that offer them, see Mohammed (2013). As seems reasonable, the main drawback is that such discounts tend to reduce the profit per unit due to the decrease in the average revenue, unless major economies of scale are achieved (Viswanathan and Wang, 2003). From the perspective of the buyer, quantity discounts mean an opportunity to reduce purchasing costs; however, buying more than needed at a specific point in time would also tend to increase inventory holding costs. Taking both effects into account, one may (simplistically) argue that quantity discounts are profitable for firms as long as the reduction in purchasing costs outweighs the increase in holding costs. However, the whole picture is more complex, given that pursuing the quantity discount has meaningful implications in the wider supply chain. For example, in line with prior discussions, large order quantities contribute to the distortion of information, thus seriously damaging the higher supply chain echelons through the Bullwhip phenomenon. Under these circumstances, the quantity discount may eventually result in increased production or transportation costs.

Due to their industrial relevance, quantity discounts have been studied in the inventory control literature over the last four decades from different angles, some of which are illustrated below by reviewing a few relevant papers in this area. Monahan (1984) developed an analytical method for establishing the optimal terms of quantity discounts from the perspective of vendors, assuming the buyer uses an EOQ model to order supplies. Weng (1995) also used an EOQ framework, through which the author explored the role of quantity discounts in the coordination of a supplier and several buyers. Li and Liu (2006) showed how quantity discounts can be used to coordinate a supplier–buyer relationship for probabilistic demands, which differs from the previous works that assumed a fixed demand. Along a different line, Meena and Sarmah (2013) used a genetic algorithm to solve the order allocation problem of a manufacturer among multiple suppliers when supply disruptions occur and quantity discounts exist. Recently, Zhao et al. (2020) studied the coordination of a fashion supply chain characterised by demand disruptions through revenue sharing contracts and quantity discounts.

This excerpt of the literature exemplifies that previous works exploring quantity discounts have provided a fair understanding of key effects caused by these pricing strategies that help professionals establish the appropriate discount in practice. However, these papers have not addressed so far the Bullwhip implications of quantity discounts in the supply chain, which play a pivotal role in many real-world industries. That is, these previous works have focused mainly on inventory-related costs, but have not considered the interplays of these costs with Bullwhip-related costs. At the same time, as discussed in the previous subsection, price considerations have been underexposed in Bullwhip studies, which have not examined the implications of quantity discounts despite their practical importance.

From this perspective, our paper attempts to shed more light on the relationship between quantity discounts and operational costs by looking at the propagation of the Bullwhip Effect in supply chains. Specifically, we measure two common metrics in the Bullwhip literature, named as Bullwhip ratio (BW) and Net Stock Amplification ratio (NSAmp), as they are symptomatic of unsteady operation in the supply chain and low efficiency in customer satisfaction, respectively. In this sense, we consider the interactions between five types of supply chain costs: (i) purchase costs; (ii) inventory holding costs; (iii) stock-out costs; (iv) capacity-related overtime costs; and (v) capacity-related opportunity costs. This analysis, which represents the contribution of our paper to the inventory control and Bullwhip literatures, allows us to derive relevant implications for professionals, which help (i) upstream supply chain managers reflect on whether or not they should offer quantity discounts to their downstream partners; and (ii) downstream supply chain managers decide on to what extent they should pursue the offer of a quantity discount made by their upstream partners.

1.3. Structure of the paper

Section 1 has positioned our research work and highlighted our contribution to the advancement of knowledge in the operations management field. To sum up, our paper contributes to the Bullwhip Effect literature by exploring the influence of quantity discounts, a popular pricing instrument that induces dynamics in supply chains that have not been investigated in enough detail so far. Also, our paper brings a new perspective to the area studying the operational implications of quantity discounts, which is based on introducing Bullwhip considerations into the analysis. Bullwhip provokes additional costs, which play a fundamental role in many supply chains, that should not be ignored in the design of quantity discounts.

The remaining of the paper has been structured as follows. Section 2 provides complete detail on the supply chain model under consideration, including the operational and economic performance metrics. Section 3 offers some preliminary analytical insights into the behaviour of the supply chain with quantity discounts. Section 4 presents our numerical analysis in a baseline scenario, carries out sensitivity analyses to explore the impact of the relevant parameters on the performance of the supply chain, and discusses the main
findings of our study. Finally, Section 5 concludes, reflects on the managerial implications of our work, and poses insightful avenues for future research.

2. Supply chain model

To understand in detail the operational dynamics induced by quantity discounts in the supply chain and the economic performance of the system, we study a single-product, serial production and distribution system formed by a manufacturer, a retailer, and the consumers. The manufacturer produces the product in response to the purchase orders issued by the retailer, who attends consumers and their product needs directly. In this section, we first present the notation of the supply chain variables and parameters. Next, we detail the sequence of events and the mathematical formulation of the supply chain model, including the underlying assumptions. Finally, we describe the performance metrics that we use for evaluating the behaviour of the supply chain, including both operational and economic indicators.

2.1. Notation

In this paper, we employ the following notation to refer to the variables that define at each moment, in a generic period $t$, the operational and financial state of the supply chain (in alphabetical order):

- $d_t$: consumer demand of the product,
- $fd_t$: forecasted demand of the product,
- $ns_t$: net stock, i.e. the on-hand inventory held by the retailer after serving the consumers,
- $o_t$: actual order quantity placed by the retailer to the manufacturer,
- $o^d_t$: product purchased by the retailer to the manufacturer at the discounted price,
- $o^r_t$: product purchased by the retailer to the manufacturer at the regular price,
- $pr_t$: product requirements, i.e. the order quantity that the retailer would go for in the absence of the quantity discount,
- $r_t$: receipts, i.e. the products received by the retailer from the manufacturer,
- $tns_t$: target net stock, representing a safety stock,
- $tw_t$: target work-in-progress, and
- $w_t$: work-in-progress, i.e. an on-order inventory of the products ordered by the retailer but not yet received.

In addition, we use the following notation for the supply chain parameters (in alphabetical order):

- $b$: unit backlog cost (an inventory-related cost),
- $dp$: discount offered by the manufacturer to the retailer in percentage terms,
- $DQ$: discount quantity, from which the economic incentive is offered,
- $h$: unit holding cost (an inventory-related cost),
- $GC$: guaranteed capacity that is available each period,
- $n$: per-unit opportunity cost (a Bullwhip-related capacity cost),
- $p$: per-unit overtime cost (a Bullwhip-related capacity cost),
- $rp$: regular price of the product,
- $Tp$: lead time, i.e. the time lag between issuing the order and receiving the product,
- $\delta$: safety parameter, i.e. the number of (additional) periods in which the safety stock protects customer service against demand uncertainties,
- $\mu$: mean of consumer demand,
- $\sigma$: standard deviation of consumer demand, and
- $\psi$: discount acceptance parameter, determining the minimum product requirements from which the quantity discount is accepted.

2.2. Mathematical model

We adopt the following sequence of events\(^5\) for describing the discrete-time operation of the production and distribution system under consideration: at the beginning of a period, the retailer receives the product corresponding to the replenishment orders placed earlier; during the course of the period, the retailer serves the demand of consumers; and at the end of the period, the retailer reviews the inventory (both net stock and work-in-progress), forecasts future demand, and issues the purchase order. At this point in time, the retailer will consider the manufacturer’s offer of a quantity discount and will decide on whether or not to order the discount quantity. This sequence of events has been modelled through a set of difference equations that is detailed below, together with the specific assumptions they entail. We will put special emphasis on the modelling of the retailer’s decision-making process around the quantity

\(^5\) This is very common in the supply chain dynamics literature, e.g. Hosoda and Disney (2006) and Ponte et al. (2017).
discount offered by the manufacturer, where the novelty of our work lies.

When the period \( t \) starts, the retailer receives the product corresponding to the order placed \( Tp + 1 \) periods ago\(^6\), as per Eq. (1). Notice that we assume that the manufacturer has always enough capacity and there is always sufficient raw material availability to produce and deliver on time, i.e., after the lead time, the quantity requested by the retailer. Then, demand occurs, which is assumed to be an independent and identically distributed (i.i.d.) random variable following a normal distribution with mean \( \mu \) and standard deviation \( \sigma \), see Eq. (2). This assumption is reasonable when demand stems from a large number of independent consumers as a consequence of the central limit theorem (Lau et al., 2013; Disney et al., 2016). Note that we also assume that \( \mu \gg \sigma \), so that the probability of negative demands is small.

\[
r_t = o_t - (Tp + 1)
\]

\[
d_t \sim N(\mu, \sigma^2)
\]

After satisfying consumer demand, both the net stock and the work-in-progress are updated. The net stock balance considers the demand (decreasing) and the receipts (increasing), as shown by Eq. (3); while the work-in-progress balance considers the last order issued (increasing) and the receipts (decreasing), as indicated by Eq. (4). Notice that \( n_s \) represents the end-of-period position of the on-hand inventory, where \( n_s > 0 \) indicates excess stock, which can be used to meet next period’s demand, while \( n_s < 0 \) reveals that stock-out has occurred. In this case, \( |n_s| \) provides the stock-out size, an unfulfilled demand (backlog) that will need to be satisfied when inventory is again available. On the other hand, it is interesting to note that, due to Eq. (1), \( Tp = 0 \) results in \( w_t = 0 \), while when \( Tp > 0 \) the work-in-progress corresponds to the sum of the last \( Tp \) orders.

\[
n_s = n_{s-1} + r_t - d_t
\]

\[
w_t = w_{t-1} + o_{t-1} - r_t
\]

The product requirements are estimated according to an order-up-to (OUT) model. We use this discrete-review replenishment policy because it is very popular in practice (Dejonckheere et al., 2003), which probably occurs as it is easy to implement and inexpensive to operate (Axsäter, 2003) and it provides good results in terms of inventory performance\(^7\) (Disney and Lambrecht, 2008). In this sense, the product requirements can be obtained as the demand forecast plus the net stock gap (i.e. target minus actual net stock) plus the work-in-progress gap (target minus actual work-in-progress), see Eq. (5).

\[
pr_t = fd_t + (ns_t - n_s) + (w_t - w)
\]

We use a static forecasting method, given by Eq. (6). Specifically, we assume the retailer has enough historical data to accurately estimate the mean demand, and uses it to forecast future demand. This is a perfectly rational decision when demand is i.i.d., as it results in a minimum mean squared error (MMSE) technique that also helps to mitigate the Bullwhip propagation in the supply chain (Disney et al., 2016). For the target net stock, we employ a safety stock model based on multiplying the demand forecast by the safety parameter \( \delta \), as per Eq. (7). This approach is also common in practice, as discussed by Hoberg et al. (2007). Meanwhile, we use the most popular target work-in-progress model (see Lin et al., 2017), where the target work-in-progress is obtained as the forecast of consumption over the lead time, see Eq. (8).

\[
fd_t = \mu
\]

\[
ns_t = \delta fd_t
\]

\[
w_t = Tp fd_t
\]

Finally, the retailer issues the purchase order, taking into account the quantity discount offered by the manufacturer. We model a ‘common-sense’ practical setting in which the retailer is only willing to accept the discount offer if the product requirements are higher than or close to the quantity from which the reduced price applies, denoted by \( DQ \). Given that the ‘close to’ concept may significantly vary among real-world retailers, we model it through the decision parameter \( \psi \), with \( 0 \leq \psi \leq 1 \), which we label the discount acceptance parameter. This parameter thus represents the willingness of the retailer to pursue the quantity discount offered by the manufacturer.

In this fashion, the retailer will only accept the discount as long as \( pr_t \geq \psi DQ \). We can illustrate this with an example. Suppose the manufacturer offers the reduced unit price from 200 units, \( DQ = 200 \). Also, assume the retailer sets a discount acceptance parameter of \( \psi = 0.9 \). If for a specific period the product requirements are \( pr_t = 175 (< \psi DQ = 180) \), the retailer will ignore the discount and order 175 units. By contrast, if the product requirements are \( pr_t = 185 (> \psi DQ = 180) \), the retailer will order the discount and order 200 units. Of course, if the product requirements are \( pr_t = 205 (> DQ = 200) \), the retailer will order 205 units (that is, this node does not only orders the minimum quantity, but exactly what it needs).

Therefore, the following rationale is assumed to reflect the decision-making process of the retailer about the quantity discount. If \( pr_t \)

\(^{6}\) \( Tp + 1 \) applies as orders are issued at the end of each period and the product is received when the period starts.

\(^{7}\) As shown by Karlin (1960), if the safety stock is appropriately adjusted, the OUT replenishment policy is optimal in terms of the sum of holding and backlog costs when both are proportional to the volume.
is higher (or equal) than $DQ$, the order equals the requirements — then, the discount will be accepted without any cost. If $pr_t$ is lower than $DQ$ but higher (or equal) than $\psi DQ$, the retailer orders the discount quantity — the discount will be accepted at the cost of ordering more than the actual needs. Lastly, if $pr_t$ is lower than $\psi DQ$, the discount will be turned down in this period, and the order again meets the exact needs of the retailer. This is described by Eq. (9).

$$o_t = \begin{cases} 
pr_t, & \text{if } pr_t < \psi DQ \\
DQ, & \text{if } \psi DQ \leq pr_t < DQ \\
pr_t, & \text{if } pr_t \geq DQ 
\end{cases}$$

(9)

Interestingly, defining the discount acceptance parameter $\psi_t$, which plays a key role in Eq. (9), will allow us to explore in detail the behaviour of the supply chain with quantity discounts, which we do in Section 4. Lastly, for evaluation of the economic metrics, it is convenient to define two additional variables. These are the number of products purchased at the regular and the discounted prices, $o'_t$ and $o''_t$, given by Eqs. (10) and (11), respectively. It can be easily verified that, of course, $o_t = o'_t + o''_t$, $\forall t$.

$$o'_t = \begin{cases} 
on, & \text{if } pr_t < \psi DQ \\
o_t, & \text{if } pr_t \geq \psi DQ 
\end{cases}$$

(10)

$$o''_t = \begin{cases} 
0, & \text{if } pr_t < \psi DQ \\
o_t, & \text{if } pr_t \geq \psi DQ 
\end{cases}$$

(11)

2.3. Operational and economic performance metrics

We quantify the operational performance of the supply chain using two common indicators that provide information of different nature but highly interrelated: the Bullwhip ratio ($BW$) and the Net Stock Amplification ratio ($NSAm$). $BW$ measures the variance of orders in relation to the variance of consumer demand, see Eq. (12); $NSAm$ assesses the variance of the net stock in relation to that of demand, see Eq. (13). As discussed by Disney and Lambrecht (2008), considering both indicators at the same time provides a rich picture of the dynamic behaviour of a specific replenishment rule; $NSAm$ reports on the inventory performance of the rule, and $BW$ is indicative of the efficiency of the production and distribution system.

$$BW = \frac{\text{var}(o_t)}{\text{var}(d_t)}$$

(12)

$$NSAm = \frac{\text{var}(ns_t)}{\text{var}(d_t)}$$

(13)

We also measure the economic performance of the system. Specifically, we consider: (i) purchase costs; (ii) inventory-related costs, including holding and stock-out costs; and (iii) Bullwhip-related capacity costs, including overtime and opportunity costs.

Purchase costs ($PC$) can be easily quantified by considering the products purchased at the regular and the discounted prices, and multiplying them by their prices, as shown in Eq. (14). To quantify the inventory-related costs ($IC$), we consider the traditional, widely used approach based on a unit holding cost of $h$ and a unit backlog costs of $b$; see e.g. Disney and Lambrecht (2008). This model is expressed by Eq. (15), where $\max(x, 0)$ is the maximum operator. To quantify the Bullwhip-related capacity costs ($CC$), we
assume that a certain guaranteed capacity $GC$ is available in each period to process the orders. If less capacity than $GC$ is needed, an opportunity cost of $n$ per unit is incurred, given that the workforce stands idle for a proportion of the period. If more capacity is required, overtime workers can be hired at a higher cost, where $p$ represents the extra cost per unit. This approach is modelled by Eq. (16). We refer readers to Disney et al. (2012) for further details behind this model.

$$PC = rpE(o_t') + rp(1 - dp)E(o_t')$$  

$$IC = hE([ns_t]^+) + bE([-ns_t]^+)$$  

$$CC = nE((GC - o_t)^+) + pE((o_t - GC)^+)$$  

In line with previous discussions, reducing $BW$ allows for a decrease in $CC$ and reducing $NSAmp$ allows for a decrease in $IC$. The rationale behind these relationships is explained in Fig. 1, where it can be seen that reducing the variability of orders (left) and inventory (right) reduces the overtime and opportunity costs (left) and the holding and stock-out costs (right), respectively. Nonetheless, it is important to highlight that to realise the improvements derived from the reduction of variabilities managers need to adjust appropriately the guaranteed capacity $GC$ (for $CC$) and the safety stock, through the safety parameter $\delta$ (for $IC$). Note, denoting in the graph the safety stock by $SS$, we obtain $SS = \delta \mu$ in our case, due to Eqs. (6) and (7).

### 3. Preliminary analytical insights

The operational dynamics of the supply chain with quantity discounts is fully characterised by Eqs. (1) to (9). As explained before, Eq. (9) describes the decision making of the retailer around the discount offered by the manufacturer. It is appropriate to note that this equation introduces a nonlinearity in the supply chain model, which has strong implications on the dynamic behaviour of production and distribution systems, as discussed by Nagatani and Helbing (2004), Wang et al. (2014), and Disney et al. (2020).

Fig. 2 represents the relationship between $o_t$ and $pr_t$ in Eq. (9) for two different values of the decision parameter $\psi$, with $\psi_1 > \psi_2$, thus providing graphical illustration of the nonlinearity under study. Comparing the $\psi_1$ and $\psi_2$ curves, we observe that as $\psi$ decreases (from $\psi = 1$) — that is, as the willingness of the retailer to pursue the quantity discount offered by the manufacturer increases —, the nonlinear nature of the supply chain is accentuated. Indeed, the nonlinearity does not exist in the specific case that $\psi = 1$, where the decision making of the retailer is not affected by the quantity discount.

This nonlinearity makes the mathematical analysis considerably more difficult, and often even intractable, unless additional assumptions are considered, such as zero lead times. For this reason, simulation has become the most popular approach to study the dynamics of nonlinear supply chain models; for instance, Chatfield and Pritchard (2013) investigated forbidden returns — a non-negative condition of orders — and Ponte et al. (2017) considered capacity constraints — an upper limit on the order quantity. In this work, we follow the same methodological approach, based on modelling and simulation techniques, to explore how quantity discounts modify the operations of supply chains and its economic consequences on the supply chain actors.

Having noted that, before the numerical, simulation-based analysis, we now discuss some preliminary analytical insights into the operational behaviour of our supply chain. To this end, Section 3.1 derives analytically the dynamics of the supply chain with $\psi = 1$, i.e. the reference linear system, against which we will explore the impact of the quantity discount. Later, Section 3.2 provides an initial understanding of the dynamic effects of the quantity discount by analysing the impulse response of the system when it faces a change in demand that moves the system from the linear to the nonlinear zone.

![Fig. 2. The relationship between the product requirements and the actual orders in our supply chain.](image-url)
3.1. Supply chain dynamics with $\psi = 1$

As we have just discussed, the interest of the retailer to get a reduced price in the product purchased to the manufacturer, indicated by $\psi < 1$, adds a nonlinearity to our model. However, for $\psi = 1$, our model behaves as a linear system, which may be interpreted as the reference system. For this reason, it is convenient to start the analysis by clarifying the dynamics of the linear system where the discount does not exist—or where it exists, but it does not alter the replenishment decisions of the retailer. In this section we use the notation $\tilde{x}$, to refer to the variables $x$ in the linear system.

To understand the system dynamics, we need to express the order ($\tilde{o}_i$) and net stock ($\tilde{n}_s$) as functions of constants and the demand ($d_i$). When $\psi = 1$, Eq. (9) simplifies to the following linear equation,

$$\tilde{o}_i = (1 + \delta + Tp)\mu - (\tilde{n}_s + \tilde{w}_i) \tag{17}$$

That is, the retailer always orders exactly what it needs. Taking this relationship into consideration, along with Eqs. (5)–(8), we can easily obtain the replenishment rule shown in Eq. (17).

Notice that $(1 + \delta + Tp)\mu$ defines a time-invariant OUT point. To understand the dynamics of the system, we can use the difference between orders placed in two consecutive periods. From Eq. (17), we see that

$$\tilde{o}_i - \tilde{o}_{i-1} = - (\tilde{n}_s + \tilde{w}_i) + (\tilde{n}_s_{i-1} + \tilde{w}_i) = (\tilde{n}_s_{i-1} - \tilde{n}_s_i) + (\tilde{w}_i - \tilde{w}_{i-1})$$

The inventory equation, Eq. (3), allows us to calculate $(\tilde{n}_s_{i-1} - \tilde{n}_s_i)$,

$$(\tilde{n}_s_{i-1} - \tilde{n}_s_i) = d_i - \tilde{r}_i.$$  

On the other hand, using the work-in-progress balance in Eq. (4), $(\tilde{w}_{i-1} - \tilde{w}_i)$ becomes

$$(\tilde{w}_{i-1} - \tilde{w}_i) = \tilde{r}_i - \tilde{o}_{i-1}.$$  

Using these two relationships, the difference between two consecutive orders can be simplified to

$$\tilde{o}_i - \tilde{o}_{i-1} = (\tilde{n}_s_{i-1} - \tilde{n}_s_i) + (\tilde{w}_{i-1} - \tilde{w}_i) = (d_i - \tilde{r}_i) + (\tilde{r}_i - \tilde{o}_{i-1}) = (d_i - \tilde{o}_{i-1})$$

Therefore, we derive the fundamental relationship of our reference linear system shown in Eq. (18).

$$\tilde{o}_i = d_i \tag{18}$$

That is, under these circumstances (in particular, OUT policy with MMSE forecasts and i.i.d. demand), the OUT policy acts as a ‘pass-on-orders’ rule, in such a way that the replenishment orders are the same as the most recent demand (see e.g. Disney and Lambrecht, 2008). In this sense, the Bullwhip phenomenon does not occur in the reference linear system, given that, using Eq. (12),

$$BW = \frac{\text{var}(o)}{\text{var}(d_i)} = \frac{\sigma^2}{\sigma^2} = 1 \tag{19}$$

Now we focus on the net stock. Using Eq. (17), the final position of the on-hand inventory can be isolated,

$$\tilde{n}_s_i = (1 + \delta + Tp)\mu - (\tilde{o}_i + \tilde{w}_i).$$

Using the relationship between receipts and past orders in Eq. (1), we can express the work-in-progress balance in Eq. (4) as follows,

$$\tilde{w}_i = \tilde{w}_{i-1} + \tilde{o}_{i-1} - \tilde{r}_i = \tilde{w}_{i-1} + \tilde{o}_{i-1} - \tilde{o}_{i-(Tp+1)} = \sum_{i=1}^{Tp} \tilde{o}_{i-i}$$

Therefore, we can consider together the sum of the order and the work-in-progress, by

$$\tilde{o}_i + \tilde{w}_i = \sum_{i=0}^{Tp} \tilde{o}_{i-i}$$

This allows us to present the net stock as a function of the past orders and, via Eq. (18), as a function of the last $(Tp + 1)$ demands. This can be seen in Eq. (20).

$$\tilde{n}_s_i = (1 + \delta + Tp)\mu - \sum_{i=0}^{Tp} \tilde{o}_{i-i} = (1 + \delta + Tp)\mu - \sum_{i=0}^{Tp} d_{i-i} \tag{20}$$

Eq. (20) allows us to derive the Net Stock Amplification ratio, defined in Eq. (13), as follows,

$$NSAmp = \frac{\text{var}(n_s)}{\text{var}(d_i)} = \frac{\sigma^2(1 + Tp)}{\sigma^2} = 1 + Tp \tag{21}$$
3.2. Dynamic analysis of the impulse response

Now we aim to understand the nonlinear discount induced by the quantity discount in the supply chain. To this end, we make use of the impulse response. Specifically, we compare the impulse response of our nonlinear supply chain with quantity discounts to that of the reference linear system. It is well known that the impulse response provides full characterisation of dynamic systems\(^8\); thus, it offers a firm understanding of the dynamic behaviour of supply chains, and hence it has been commonly employed in the literature (e.g. Disney and Towill, 2003; Dominguez et al., 2014; Gaalman et al., 2019).

Unless unreasonable assumptions are made, our system will often be operating in the linear region of Fig. 2, due to \(pr_t < \psi DQ\). At some specific points in time, in particular when \(pr_t \geq \psi DQ\) and \(pr_t < DQ\), it will enter the nonlinear region of Fig. 2. When the supply chain is operating as a linear system, \(pr_t = \psi DQ\) and the arrival of a demand, \(d_t\), that verifies \(\psi DQ \leq d_t < DQ\) is the condition that will make it enter the nonlinear region. To illustrate the short-term dynamics induced by the discount, we then need to generate a demand impulse so that, \(\forall t \neq 0\), \(d_t < \psi DQ\) and, for \(t = 0\), \(\psi DQ \leq d_t < DQ\). By way of illustration, we use \(DQ = 140\) and \(\psi = 0.75\) (\(\psi DQ = 105\)) and we consider the impulse function, \(d_t = \{100, \forall t \neq 0; 110, \text{ if } t = 0\}\). Also, we use a lead time of \(T_p = 4\) periods.

Fig. 3 shows the impulse response of the linear and nonlinear systems.

Consider that during a short number of consecutive periods, the supply chain is operating in the linear region, given that the last demands were lower than \(\psi DQ\), i.e. \(d_{t-1}, d_{t-2} \cdots < \psi DQ\). Fig. 3 shows that every time \(t\) the supply chain enters into the nonlinear region, due to \(\psi DQ \leq d_t < DQ\), the following sequential behaviour emerges, establishing key differences between the linear and nonlinear systems:

i. In \(t, \psi DQ \leq d_t < DQ\) leads to \(\psi DQ \leq pr_t < DQ\), which according to Eq. (9) makes that \(o_t = DQ\). This creates a (first) gap between the linear orders, \(\tilde{d}_t\), and the nonlinear orders, \(o_t = DQ\), which we denote by \(\Delta_t = DQ - d_t\). In Fig. 3, \(\Delta_{t-0} = 140 - 110 = 30\).

ii. In \(t + 1\), the work-in-progress increases in the nonlinear system by \(\Delta_t\) units due to the last order being larger than the linear one. That is, as per Eq. (4), \(w_{t+1} = w_{t} + \Delta_t\). Using Eq. (5), this provokes a reduction in the product requirements, \(pr_{t+1} = pr_{t+1} - \Delta_t\). Due to Eq. (9), this will make that \(o_{t+1} = \tilde{d}_{t+1} - \Delta_t\), except if the new demand does not allow for it.

iii. From \(t + 2\), the nonlinear system behaves as the linear system would, given that the excess of products ordered in \(t\) has been compensated in \(t + 1\). Note that \(o_{t+1} = \tilde{d}_{t+1} - \Delta_t\) makes that \(w_{t+2} = (w_{t+2} + \Delta_t) - \Delta_t = w_{t+2}\). Notice that the difference for the other variables shown in Fig. 3 is also zero. However, this only occurs up to \(t + T_p\), inclusive (\(t + 4\), in Fig. 3).

iv. In \(t + T_p + 1\), the product associated to the order issued in \(t\) is received at the on-hand stock site. That is, is via Eq. (1), \(r_{t+T_p+1} = r_{t+T_p+1} + \Delta_t\). Therefore, as per the net stock and work-in-progress balances in Eqs. (3), \(ns_{t+T_p+1} = ns_{t+T_p+1} + \Delta_t\) and \(w_{t+T_p+1} = w_{t+T_p+1} - \Delta_t\).

v. In \(t + T_p + 2\), the ‘compensation order’ of \(t + 1\) is received, i.e. \(r_{t+T_p+2} = r_{t+T_p+2} - \Delta_t\). See Eq. (1). The net stock and work-in-progress return to their linear states; \(ns_{t+T_p+2} = (ns_{t+T_p+2} + \Delta_t) - \Delta_t = ns_{t+T_p+2}\), \(w_{t+T_p+2} = (w_{t+T_p+2} - \Delta_t) + \Delta_t = w_{t+T_p+2}\) (Eqs. 3–4). Thus, the impact of the discount used in \(t\) on the system dynamics ends in \(t + T_p + 2\).

Steps (i) to (v) occur repeatedly in our supply chain. When the minimum quantity from which the discount is accepted is sufficiently higher than the mean demand, i.e. \(\psi DQ \gg \mu\), the probability of having two consecutive demands that are higher than \(\psi DQ\) is negligible. For example, for \(\mu = 100\), \(\sigma = 25\), \(DQ = 180\) and, \(\psi = 75\%\), we obtain \(P(d_t > \psi DQ)(d_{t-1} > \psi DQ) < 1\%). In this case, \(o_t = DQ\) will (nearly) always be compensated in \(t + 1\), and the effects of the discount do not overlap in time. When the previous condition does not hold, the effects of different discounts will overlap in time.

In any case and for a long time horizon, as in Fig. 3, a variable in the nonlinear system \(x_t\) can be expressed as the sum of the linear variable, \(\tilde{x}_t\), and the difference between them, denoted by \(\tilde{x}_t\), i.e. \(x_t = \tilde{x}_t + \tilde{x}_t\). This decomposition allows us to understand the behaviour of the nonlinear variables. We continue the analysis by focusing on the order \((\tilde{o}_t)\) and net stock \((\tilde{n}_s)\), which explain the performance metrics of our system.

For the orders, \(o_t = \tilde{o}_t + \tilde{o}_t\). As \(\tilde{o}_t\) and \(\tilde{o}_t\) are not independent, but correlated variables, we obtain that \(\text{var}(o_t) = \text{var}(\tilde{o}_t) + \text{var}(\tilde{o}_t) + 2\text{cov}(\tilde{o}_t, \tilde{o}_t)\). We have previously shown that \(\text{var}(\tilde{o}_t) = \sigma^2\), see Eq. (19). By the variance definition, \(\text{var}(\tilde{o}_t) \geq 0\); indeed, considering that \(E(\tilde{o}_t) = 0\) due to the compensation effect in steps (i) and (ii), \(\text{var}(\tilde{o}_t) > 0\), except if \(\tilde{o}_t = 0\) (which only happens if the nonlinear region does not exist in Eq. (9), or the probability of entering it is null). In addition, \(\tilde{o}_t\) and \(\tilde{o}_t\) are correlated variables, see Fig. 3. Under the normal assumption that \(DQ > \mu, \tilde{o}_t > 0\) occurs when \(\tilde{o}_t = d_t\) takes high values, that is, \(\text{cov}(\tilde{o}_t, \tilde{o}_t) > 0\). Thus, \(\text{var}(\tilde{o}_t) > \text{var}(\tilde{o}_t)\). In this way, the discount tend to increase the Bullwhip Effect, as also suggested by the increase of order variability in Fig. 3.

Nonetheless, it is also interesting to note that, if \(DQ < \mu\), the nonlinear behaviour generally occurs when \(\tilde{o}_t = d_t\) is low (for high values of \(\tilde{o}_t = d_t\), the system operates again in the linear region, see Fig. 2). In this case, \(\text{cov}(\tilde{o}_t, \tilde{o}_t) < 0\). While it may not be reasonable for the manufacturer to offer the discount under such circumstances, this may lead the supply chain to benefit from a reduced BW ratio.

We now perform a similar analysis for the net stock. In this case, \(ns_t = \tilde{n}_s + \tilde{n}_s\), which leads to \(\text{var}(ns_t) = \text{var}(\tilde{n}_s) + \text{var}(\tilde{n}_s) + \text{cov}(\tilde{n}_s, \tilde{n}_s)\)

---

\(^8\) The impulse response is widely used in the study of linear dynamic systems. Its analytical importance lies in the fact that the impulse response is the inverse Laplace transform of the transfer function that relates the outputs under consideration with the relevant input(s); see e.g. Yang et al. (1995). For this reason, it can be interpreted as a graphical representation of the function that relates the input(s) and the output(s) of the system.
As per Eq. (21), \( \text{var}(\tilde{n}_t) = \sigma^2(1 + Tp) \). Again, due to the definition of variance, \( \text{var}(\tilde{n}_t) > 0 \), except if the discount does not affect the supply chain operation. Lastly, note that the differences between the nonlinear and linear net stocks, \( \tilde{n}_t > 0 \), are provoked by the demand \((Tp + 1)\) periods before, i.e. \( d_{t-(Tp+1)} \). In contrast, \( n_t \) depends on demands from \( d_t \) up to \( d_{t-(Tp+1)} \), in line with Eq. (20). Under these circumstances, the covariance term does not play a meaningful role in the previous relationship, and \( \text{var}(n_t) > \text{var}(\tilde{n}_t) \). Thus, the discount also tends to increase the \( \text{NSAmp} \) ratio, which is aligned with the analysis of net stocks in Fig. 3.

All in all, we conclude from our preliminary analytical study that \( BW \) and \( \text{NSAmp} \) in the nonlinear supply chain will tend to be higher than that in the linear one due to the impact of the retailer’s pursuit of the discount offered by the factory. The impulse response has allowed us to understand why the deterioration of the dynamics occurs in the inventory system. However, this does not necessarily apply in the unusual case that \( DQ < \mu \), where the supply chain may benefit from an improved order dynamics.

4. The supply chain dynamics of the quantity discount

We now study in depth the operational and economic performance of the production and distribution system with quantity discounts when exposed to the stochastic conditions defined. To analyse the effects of the quantity discount, the discount acceptance parameter, \( \psi \), plays a key role, as it models the retailer’s willingness to pursue the quantity discount when making an order. We define a baseline scenario represented by the values of the parameters indicated below, and justified as follows:

- We consider a mean consumer demand of \( \mu = 100 \) (units) and a standard deviation of \( \sigma = 25 \) (units). This implies a c.v. of \( \sigma / \mu = 25\% \), which is within the typical range of variation of the demand time series faced by retailers, [15\%, 50\%], observed by Dejonckheere et al. (2003).
- We assume that there is a lead time of two periods, \( Tp = 2 \) (periods), i.e. the product is received by the retailer two periods after it is sent by the manufacturer. This is a typical, frequently used value of the lead time in supply chains studies, such as e.g. Ciancimino et al. (2012).
- We consider that the regular price of the product is \( rp = 10 \) (€/unit). Also, we consider that the manufacturer offers a discount of \( dp = 10\% \) —resulting in a discounted price of \( rp(1 - dp) = 9 \) (€/unit)— when at least 140 units are ordered, that is, \( DQ = 140 \) (units). This percentage (10\%) may be interpreted as a reasonable quantity discount found in practice that authors have used in prior literature, e.g. Lightfoot (2019).
- We select \( b = 2 \) (€/unit) and \( h = 1 \) (€/unit), representing a traditional scenario in which backlog is more expensive than holding inventory. This is by far the most common case in practice —otherwise, real-world companies would not need to
operate with safety stocks—; for example, \( b = 2h \) is generally used in the well-known Beer Game, see Goodwin and Franklin (1994).

- We select \( p = 2 \) (€/unit) and \( n = 1 \) (€/unit), representing a practical setting in which overtime is more costly than the opportunity cost of being idle. Other studies have also made this assumption, such as Ponte et al. (2017), who also use \( p = 2n \). In real-world practice, it may also occur that \( n > p \), e.g. if large investments have been made, which will be explored later.
- Due to the lack of precise knowledge of the nonlinear system with quantity discounts, we assume that the retailer has optimised the performance of the system in the linear setting by using Eqs. (A.1) and (A.2) in Appendix A to determine the appropriate safety parameter and guaranteed capacity. In this baseline scenario, this results in \( \delta^* = 0.1865 \) and \( GC^* = 110.77 \) (units).

In Section 4.1, we analyse the long-term behaviour of the system in the baseline scenario. We aim to gain general insights into the dynamic consequences of the manufacturer’s offer of a quantity discount, detecting the key cause-effect relationships motivated by the decision parameter \( \psi \). In Section 4.2, we perform several sensitivity studies with the aim of exploring the impact of the other parameters on the performance of the system. This allows us to gain a greater understanding of the operational and economic effects of quantity discounts in different types of real-world supply chains.

Our analysis is based on simulating the long-term response of the stochastic system. This methodological perspective is aligned with many works that claim that simulation techniques can be a very effective and efficient approach to investigate complex, nonlinear logistics systems (e.g. Iannoni and Morabito, 2006; Chatfield and Pritchard, 2013; Matopoulos et al., 2016; Dominguez et al., 2018; Ponte et al., 2018). To this end, we have implemented the difference equations defining our supply chain model in MATLAB R2019b. We have carried out simulation runs of 400,000 periods to ensure the repetitiveness and consistency of the results. In each case, we have simulated the supply chain for 41 values of \( \psi \) in the interval \( 0 \leq \psi \leq 1 \), i.e. \( \psi = \{0.0.025, 0.050, \ldots, 1\} \), to better perceive the impact of this parameter.

### 4.1. Analysis of the baseline scenario

Fig. 4 displays the operational performance of the supply chain with quantity discounts by showing the relationship of the BW and NSAmp metrics with the discount acceptance parameter, \( \psi \). To interpret the curves, it is important to recall that \( \psi = 1 \) defines the behaviour of the linear system —where the dynamics are not affected by the quantity discount. As it is known that the linear system is characterised by \( BW = 1 \) and \( NSAmp = 1 + Tp = 3 \) (see Section 3.1), we can observe that for \( \psi = 1 \) our nonlinear system behaves as expected. The main finding of Fig. 4 is that decreasing \( \psi \) provokes a sharp increase of \( BW \) and \( NSAmp \), which is in line with the insights derived in Section 3.2 from the impulse response analysis. That is, pursuing the quantity discount significantly increases the volatility in the operation of the retailer, which would also be transmitted upstream in the supply chain.

A more detailed inspection of Fig. 4 reveals that, in this scenario, the increase of variability is more accentuated for orders (BW) than for inventories (NSAmp). For example, for \( \psi = 0.8 \), we see that NSAmp has only marginally increased (from NSAmp = 3) to NSAmp \( \approx 3.1 \), while BW has strongly risen (from BW = 1) up to BW \( \approx 1.5 \). This result points to the more negative impact of the quantity discount on order variability (which has implications in terms of production smoothness and transportation efficiency, among others) in comparison with inventory variability (with implications on the supply chain ability to satisfy consumer demand in a cost-efficient way).

In this way, our analysis confirms that pricing considerations take a key role in the propagation of the Bullwhip Effect in supply chains; in particular, quantity discounts ‘force’ echelons to unintentionally amplify the variability in the system. Note that the more willingness of the supply chain nodes to pursue the discount, the higher amplification of the variabilities in the system.

Thus, the operational analysis suggests that pursuing the quantity discount will tend to increase both the inventory-related costs (because of the increased NSAmp) and, more noticeably, the Bullwhip-related capacity costs (due to the increase in BW). To analyse the economics of the system, Fig. 5 provides information on the impact of \( \psi \) on the inventory-related costs (left) and the capacity costs (right).

The Bullwhip-related costs analysis of Fig. 5 is well aligned with the BW curve in Fig. 4. In general terms, we observe that reducing \( \psi \) significantly increases the capacity costs. For instance, when \( \psi \approx 0.625 \), the capacity costs are the double than in to the linear scenario (\( \psi = 1 \)). This occurs because decreasing \( \psi \) increases both the overtime and the opportunity costs in our supply chain; although it can be noted that the increase in overtime costs is more significant in this baseline scenario. This finding can be easily explained: to pursue the quantity discount, the retailer will need to resort frequently to overtime, given that \( DQ > GC^* \). All in all, we conclude that the willingness to pursue the discount provokes a dramatic increase of Bullwhip-related costs in the supply chain, unless the discount is only accepted for product requirements that are quite close to the discount quantity, say, \( \psi > 0.9 \).

In relation to inventory costs, it is interesting to note that the quantity discount causes a reduction in backlog costs. In other words, pursuing the discount helps the company achieve higher customer service levels. This also seems reasonable; ordering more than the actual needs to achieve the discounted price protects the supply chain against stock-outs in the near future. However, it is relevant to note that this comes at the expense of a stronger increase in holding costs, which makes that the total inventory-related costs increase as \( \psi \) reduces despite the improvement in the service level. In line with the BW vs. NSAmp analyses, the impact of \( \psi \) on inventory-related costs is lower in relative terms to that on capacity costs. Indeed, as long as \( \psi > 0.8 \), inventory costs are not severely affected by the discount.

Overall, the quantity discount increases the inventory- and Bullwhip-related costs of the retailer. Even so, it may be individually
Fig. 4. Operational metrics in the baseline system.

Fig. 5. Inventory- and Bullwhip-related costs in the baseline system.

Fig. 6. Purchase and total costs in the baseline system.
rational for this node to pursue the discount due to the reduction in purchase costs. Therefore, to have a complete picture of the effects of the discount, we also need to consider purchase costs. These are represented in Fig. 6. For the sake of completeness, this figure also includes a curve that expresses the total cost assumed by the retailer—which is the sum of Bullwhip-related, inventory-related, and purchase costs—as a function of the discount acceptance parameter, $\psi$.

First, Fig. 6 shows that decreasing $\psi$ allows for a significant reduction of purchase costs. Note that the slope of the curve is stronger for high $\psi$; therefore, the retailer may significantly benefit from reducing the value of the decision parameter $\psi$ from the initial $\psi = 1$. Also, as we discussed before, the increase in Bullwhip- and inventory-related costs is moderately small for high values of $\psi$; therefore, it may be profitable to pursue the discount, i.e. to use $\psi < 1$, despite its negative impact on the operation of the supply chain. In contrast, for low $\psi$, a further decrease of this parameter provokes a relatively small decrease of purchase costs that comes at the expense of a large increase in Bullwhip- and inventory-related costs. Thus, it would not generally be cost-efficient to set $\psi$ close to $\psi = 0$.

The previous arguments suggest the existence of an optimal discount acceptance parameter, between $\psi = 0$ and $\psi = 1$, which we denote by $\psi^*$. This existence is confirmed by the total cost curve in Fig. 6, which adopts a convex function form. In this sense, professionals can optimise the performance of their production and distribution systems with quantity discounts by appropriately regulating the decision parameter $\psi$. In the baseline system, Fig. 6 reveals that the optimal value is $\psi^* \approx 0.6$. In the next subsection, we will investigate the effect of the different parameters on the dynamics of the system with quantity discounts across with their impact on the optimal calibration of $\psi$.

4.2. Sensitivity analysis of the supply chain parameters

First, we address the operational parameters: the standard deviation of consumer demand ($\sigma$), the lead time ($Tp$), and the discount quantity ($DQ$). Once we have clarified the links of $BW$ and $NSAmp$ with the capacity and inventory-related costs, respectively, our analysis focuses on the $BW$ and $NSAmp$ metrics along with the total costs faced by the retailer.

Fig. 7 represents the indicators as functions of $\psi$ in three scenarios of demand variability: low ($\sigma = 10$, c.v. = 10%), medium ($\sigma = 25$, c.v. = 25%), i.e. the baseline scenario), and high ($\sigma = 40$, c.v. = 40%). All three curves tend to $BW = 1$ and $NSAmp = 1 + Tp = 3$ as $\psi$ approaches $\psi = 1$, given that these metrics do not depend on $\sigma$ in the linear system. In this fashion, the curves diverge as $\psi$ decreases; that is, the impact of demand variability is stronger for lower values of the discount acceptance parameter. It is interesting to note that the negative effects of reducing $\psi$ on $BW$ and $NSAmp$ are attenuated as $\sigma$ grows, which occurs because these metrics are relative to the variability of demand.1 For this reason, the discount acceptance parameter that minimises the total costs, $\psi^*$, is a decreasing function of $\sigma$: for $\sigma = (10, 25, 40)$, $\psi^* \approx (0.75, 0.60, 0.45)$, respectively. That is, when demand variability is low, managers should use higher values of $\psi$; otherwise, the system will suffer from a deteriorated dynamics.

Fig. 8 shows the same information for three lead-time scenarios defined by $Tp = \{1, 2, 4\}$, where $Tp = 2$ represents the previously analysed baseline scenario. Looking at the $BW$ and $NSAmp$ curves, we find that the impact of $Tp$ in the nonlinear system is the same as, or at least very similar to that, in the linear system, i.e. the lead-time effects do not depend noticeably on $\psi$. Note that $BW$ is not affected by $Tp$, while $NSAmp$ increases by $Tp$ units as a result of the lead-time effects; please recall that $BW = 1$ and $NSAmp = 1 + Tp$ in the linear system due to use of MMSE forecasts under i.i.d. demand (see Section 3.1). Finally, the total cost curve shows how reducing $Tp$ allows for a decrease in the costs of the retailer. In this regard, detailed inspection of the graph suggests that $\psi^*$ decreases as $Tp$ increases (specifically, for $Tp = \{1, 2, 4\}$, $\psi^* \approx (0.65, 0.60, 0.55)$, respectively); nonetheless, the impact of the lead time on the optimal selection of the discount acceptance parameter is lower than for the rest of parameters.

Fig. 9 considers the implications of the discount quantity. Together with the baseline scenario defined by $DQ = 140$, we now use $DQ = 120$ to represent the case in which the discount quantity is relatively close to the mean demand, and $DQ = 180$ to illustrate the opposite case. Here, the $BW$ and $NSAmp$ curves reveal that reducing $\psi$ has more damaging consequences on the dynamics of the supply chain for high $DQ$.\textsuperscript{10} This occurs because, when $DQ$ is close to $\mu$, the discount can be often accepted at a relatively low operational cost; however, when $DQ \gg \mu$, pursuing the discount significantly reduces the smoothness of supply chain operation. Under such circumstances, setting the discount acceptance parameter to low values is more costly in the latter case (i.e. when $DQ \gg \mu$). This explains why, in the total cost curve, the optimal $\psi$ proves to be an increasing function of $DQ$; notice that, for $DQ = 120$, $\psi^* \approx 0.5$; for $DQ = 140$, $\psi^* \approx 0.6$; and for $DQ = 180$, $\psi^* \approx 0.7$.

It may also be convenient to consider the particular scenario where $DQ < \mu$. Thereby, Fig. 9 also displays the supply chain with $DQ = 80$. In line with the preliminary insights discussed in Section 3.2, we now observe that $BW$ reduces as $\psi$ decreases; interestingly, in this case the discount offer has a positive impact on the order dynamics. The same positive effect cannot be observed in terms of $NSAmp$, but it can be noted that the negative impact of the discount on this metric is lower than for $DQ > \mu$. In such cases, the retailer should opt for unusually low values of the discount acceptance parameter.

---

\textsuperscript{10} When $\psi$ decreases, the operational variability becomes heavily influenced by the discount, and hence is more insensitive to $\sigma$. Thus, for a given $\sigma$, the differences in the numerator of the $BW$ and $NSAmp$ metrics decrease when $\psi$ is reduced (as the discount governs the dynamic of the response); while the denominator is not affected ($\text{var}(d_t) = \sigma^2$); see Eqs. (12) and (13).

\textsuperscript{11} As discussed, the general rule is that, for a given $\psi$, the increase of $DQ(> \mu)$ results in an increased $BW$ and $NSAmp$. However, it is interesting to note that when $\psi > 0.8$, the lowest $BW$ and $NSAmp$ is obtained for the highest $DQ(> \mu)$. This occurs given that, when $\psi DQ > \mu + 2\sigma$ (i.e. when $Q > \frac{\mu + 2\sigma}{\psi}$), the product requirements, $pr_t$, are nearly always lower than the minimum quantity to pursue the discount, $\psi DQ$. This implies that the discount is rarely pursued by the retailer, and hence our nonlinear system behaves very close to the linear one.
Now we consider the economic parameters; specifically, the discount offered by the manufacturer in percentage terms, $dp$, and the four unit costs, $\{b, h, p, n\}$. As the curves of the operational metrics do not depend on them, our sensitivity analysis will now directly look at the total cost curve.

The impact of the discount, $dp$, is represented graphically in Fig. 10, using $dp = \{5\%, 10\%, 20\%\}$. We can see that increasing $dp$ has positive consequences on the economics of the retailer. Also, the figure reveals that the optimal $\psi$ is very sensitive to $dp$. As discussed, in the baseline scenario ($dp = 10\%$), $\psi^* \approx 0.6$. For higher discounts, the retailer should react by setting $\psi$ to very low values with the aim of taking frequent advantage of the discount; note that $dp = 20\%$ results in $\psi^* \approx 0.2$. That is, as $dp$ grows, it seems reasonable to increase the willingness to order $DQ$ in an attempt to minimise purchase costs. In contrast, for lower $dp$, $\psi$ should take higher values; e.g. $dp = 5\%$ results in $\psi^* \approx 0.85$.

At this point, it is convenient to underline that, while the relationship between the operational metrics ($BW$ and $NSAmp$) and the discount acceptance parameter ($\psi$) does not directly depend on the percent discount ($dp$), it can be expected that $dp$ impacts the operational performance of the system. Note that $BW$ and $NSAmp$ depend on $\psi$, which in turn should be established by taking $dp$ into account. In this regard, as we have just discussed, supply chain managers should opt for a lower $\psi$ as $dp$ grows. Therefore, for high $dp$, the dynamics of the supply chain will be more deteriorated, as per Fig. 4.

Lastly, we study the unit costs, $\{b, h, p, n\}$, which in the baseline case adopted the values $\{2, 1, 2, 1\}$. The top graphs of Fig. 11 modify the unit costs under the same $b/h$ and $p/n$ ratios. The left one considers $b$ and $h$. When they grow — which refers to scenarios in which inventories are especially costly, both in terms of holding extra stock and stock-out occurrence — $\psi$ should take higher values. The right graph, showing $p$ and $n$, reveals similar insights; we can see that as the cost of order variability increases (for example, when capacity is expensive and inflexible), $\psi^*$ grows. Notice, when $\{b, h\}$ increase from $\{2, 1\}$ to $\{4, 2\}$, $\psi^*$ grows from 0.6 to 0.7 (approx.); if this happens in $\{p, n\}$, $\psi^*$ becomes 0.85 (approx.). That is, when the unit costs of holding inventory and backlog or (and) opportunity

![Fig. 7. The impact of the variability of consumer demand.](image-url)
and overtime are particularly expensive, managers should avoid pursuing the quantity discounts from low orders quantities. Under such circumstances, the operational deterioration of the supply chain response caused by the discount results in a significantly reduced economic performance.

The bottom graphs of Fig. 11 vary the unit costs, keeping \( b + h \) (left) and \( p + n \) (right) constant. First, we focus on \( \{b, h\} \). When the relative importance of backlog costs increases, \( \psi^* \) decreases; e.g. from \( \{b, h\} = \{2, 1\} \) to \( \{2.5, 0.5\} \), \( \psi^* \) reduces from 0.6 to 0.45 (approx.). Equivalently, when holding costs are higher than backlog costs, \( \psi^* \) grows; note that for \( \{b, h\} = \{1.2\} \), \( \psi^* = 0.65 \). Thus, only if backlog is significantly more expensive than holding, managers should set \( \psi \) to relatively low values. To understand this finding, we may need to refer back to Fig. 5, which showed that reducing \( \psi \) contributes to decreasing backlog costs. Finally, we study \( \{p, n\} \). We observe that as the \( p/n \) ratio is reduced, \( \psi^* \) increases. Therefore, in practical scenarios in which overtime is much more costly than regular working time, managers should show a higher willingness to pursue the quantity discount. For instance, inspection of the graph reveals that \( p/n = 0.5 \) leads to \( \psi^* \approx 0.7 \), while \( p/n = 6 \) results in \( \psi^* \approx 0.45 \).

To sum up, Table 1 provides an overview of the effect of the main characteristic parameters of our production and distribution system on the optimal adjustment of the discount acceptance parameter, \( \psi^* \). Interpreting the findings of this table in their specific contexts should provide supply chain managers with valuable information on to what extent they should pursue offers of quantity discounts.

5. Conclusions

Offering quantity discounts is a common pricing strategy in many industries, through which upstream supply chain members entice their downstream partners to purchase in larger quantities. These pricing mechanisms give rise to meaningful interactions between the
Fig. 9. The impact of the discount quantity.

Fig. 10. The impact of the percent discount.
marketing and the operations functions of organisations that have been considered to certain extent in previous works. However, the implications of quantity discounts on the dynamics of supply chains have remained largely unexplored, mainly due to the complexity of bringing pricing considerations into the operational analysis.

In this paper, we investigate how the manufacturer’s offer of a quantity discount and the retailer’s willingness to pursue it affect the operational and economic performance of supply chains. Importantly, we provide evidence of how the quantity discount tends to deteriorate the dynamics of these systems by contributing to the Bullwhip Effect propagation. Issuing orders that do not match the actual requirements of the organisation may greatly distort the transmission of information along supply chains, which negatively impacts not only order variability but also inventory performance. In this sense, we reveal that the quantity discount also hinders the satisfaction of customer demand in a cost-efficient manner. As an exception, we observe that discounts may alleviate the Bullwhip phenomenon when the discount quantity is lower than the mean demand. However, offering the discount under these conditions might not be reasonable for the manufacturer even if that improved the dynamics of the system.

Although our study also has valuable implications for offerors of quantity discounts (discussed in the next subsection), we adopt the perspective of a retailer that considers to what extent to pursue the discount offered by the supplier. We observe that, even though the discount increases the inventory- and capacity-related costs of the retailer, it may be reasonable for this node to pursue the discount, given that this would decrease the purchase costs. Importantly, our analysis shows the existence of a U-shaped, convex function that

| Table 1 |
|---------|
| The effects of the different factors on the optimal discount acceptance parameter. |
| Factor | Effect on $\psi^*$ | Factor | Effect on $\psi^*$ |
|--------|-----------------|--------|-----------------|
| Variability of consumer demand | ↓ | Importance of inventory-related costs | ↑ |
| Lead time | ↓ | Relative weight of backlog to holding costs | ↓ |
| Discount quantity relative to mean demand | ↑ | Importance of capacity-related costs | ↑ |
| Discount percentage | ↓ | Relative weight of overtime to regular time costs | ↓ |

Note: The table shows the effect of increasing the strength of the different factors on the optimal value of $\psi$. |
relates the total cost with the retailer’s willingness to pursue the discount, which we model through the discount acceptance parameter. In this fashion, practitioners can optimise the economic performance of their system by setting appropriately the decision parameter $\psi$. In addition, we explore the effect of the rest of operational and economic parameters on the retailer’s optimal discount acceptance parameter, leading to practical managerial insights on which we elaborate below.

5.1. Implications for professionals

The reported implications of quantity discounts on the Bullwhip Effect should be carefully taken into consideration by sellers that consider the option of offering them to their customers. For instance, this applies to manufacturers. While quantity discounts have the potential to help manufacturers increase sales revenue and achieve economies of scale, under some circumstances they will make production more unstable. In the words of the Lean management paradigm, quantity discounts may contribute to the costly wastes of mura and muri, closely related to the Bullwhip phenomenon.

In the light of these findings, our paper highlights the need for considering the interdependencies among the various business functions when making both strategical and tactical decisions. Otherwise, optimising some processes may occur at the expense of creating meaningful inefficiencies from other perspectives, whose actual impact on business performance tends to be underestimated. This is especially relevant in the management of supply chains, due to their complex and dynamic nature. Thus, our paper should encourage upstream supply chain echelons, such as manufacturers, to explore the indirect and non-immediate consequences of their pricing decisions, as they may result in a significant deterioration of the operational dynamics of their supply chains.

All in all, the unintended operational consequences that quantity discounts create in the supply chain may discourage some manufacturers from offering such discounts to their customers. Others, however, may still opt for the offer of a discount for a wide variety of reasons. In such cases, we would strongly recommend manufacturers to promote the implementation of collaborative practices in their supply chains, as this would significantly alleviate the negative dynamics induced by the discount; for example, those practices based on the synchronisation of supply chain operations, see Ciancimino et al. (2012). When implementing collaborative practices, the integrative framework for supply chain collaboration developed by Simatupang and Sridharan (2005) may be of special interest for professionals.

Now we focus on the perspective of buyers that need to consider whether to modify their replenishment decisions due to the discount offer. Our work reveals that it is fundamental to compare the benefits of quantity discounts, mainly in the form of a reduced purchase price, to the additional costs derived from operating with a higher variability than needed, not only in terms of inventory-related costs but also in terms of Bullwhip-related costs. Our results highlight that there is a significant economic benefit in determining the approximate form of their convex curves that relate the total cost with the discount acceptance parameter, $\psi$ —or, equivalently, the minimum quantity from which the discount is accepted.

For obvious reasons, our paper is unable to provide practitioners with the exact form of their curves. However, our sensitivity analysis offers valuable clues for managers on how to accurately ‘tune’ their decision parameter $\psi$. In this regard, uncertainty in customer demand plays a key role. Interestingly, managers facing highly uncertain demands should calibrate $\psi$ to lower values. Also, while the lead-time effects are less strong, retailers operating in supply chains with long lead times should set $\psi$ to lower values. Taking both into account, it is interesting to note that inventory managers that need to deal with tougher environmental conditions (characterised by high demand uncertainty and long lead times) should show a higher willingness to pursue the quantity discount offered by their suppliers.

The discount percentage as well as the ratio of the discount quantity to the mean demand also have significant implications on the optimal discount acceptance parameter, $\psi$. Specifically, $\psi^*$ decreases as the discount percentage increases and this ratio reduces. In such cases, it is individually rational to alter the operational dynamics of the supply chain with the aim of taking advantage of the reduced price. In addition, when inventory- and/or capacity-related costs are especially costly for the retailer, managers should prevent from being especially keen to pursue the discount quantity. However, when the relative weight of backlog to holding costs is low (e.g. where storage space is especially costly) and the relative weight of overtime to regular time costs is low (e.g. when overtime working is affordable), retailers should not be reluctant to pursue the discount from relatively low product requirements.

5.2. Avenues for future research

The Bullwhip Effect should pay particular attention to the effect of price considerations in this upcoming decade. Despite Lee et al. (1997a,b) pointed out that pricing strategies are one of the major causes of Bullwhip, very little has been investigated so far on how they add to the propagation of this harmful phenomenon, as highlighted by modern reviews of the literature (Bhattacharya and Bandyopadhyay, 2011; Wang and Disney, 2016). It thus becomes necessary to explore in detail how popular pricing strategies impact supply chain dynamics, such as mark-up pricing, skimming, and penetration pricing. In a related way, it would be interesting to study how the psychology of consumption and pricing (Gourville and Soman, 2002) affects the behavioural causes of the Bullwhip Effect.

Also, the Bullwhip discipline in the 2020 decade will need to be populated with more studies delving into the dynamics of closed-loop supply chains. As sustainability acquires a critical importance, the supply chain structure is evolving from linear models, such as the one studied in this paper, to closed-loop variants (e.g. Xiao et al., 2020). It thus becomes necessary to explore the effect of pricing mechanisms on the performance of these new supply chains that incorporate additional sources of complexity (Goltsos et al., 2019). Note that important interplays emerge here between the pricing of new and remanufactured products (Atasu et al., 2010) that need to be investigated from a dynamic perspective.

Finally, we point out to the emerging concept of organisational resilience, which has gained even more importance during the
current COVID-19 pandemic (Ivanov, 2020). Disruptions provoke a Bullwhip-like phenomenon that is labelled as the Ripple Effect of supply chains (Ivanov et al., 2014). Investigating the interactions between pricing decisions and disruption propagation is also an area of research worth pursuing, which would yield relevant implications for supply chain professionals.

CRediT authorship contribution statement

Borja Ponte: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization. Julio Puche: Methodology, Software, Validation, Writing - review & editing, Visualization. Rafael Rosillo: Formal analysis, Investigation, Writing - review & editing. David Fuente: Conceptualization, Resources, Writing - review & editing, Supervision.

Appendix A. – Optimisation of the linear system

When backlog and holding costs are proportional to the volume, the optimal safety stock, \(SS^*\), in the linear system can be expressed as

\[
SS^* = \sigma_o \Phi^{-1} \left[ \frac{b}{b + h} \right]
\]

where \(\Phi^{-1}\) is the inverse of the cumulative distribution function of the standard normal distribution and \(\sigma_o\) is the standard deviation of the net stock levels over time (e.g. Disney and Lambrecht, 2008). To adapt this equation to the specific context of our paper, we can use that in our linear system \(NSAmp = \frac{\var(\sigma _{n s})}{\var(\sigma _{n s}^o)} = 1 + Tp\), see Section 3.1. Note, the Net Stock Amplification ratio can also be presented as \(NSAmp = \frac{\sigma}{\sigma^o}\). Therefore, \(\sigma_o = \sigma \sqrt{1 + Tp}\). Finally, we take into consideration that the relationship between the safety stock, \(SS\), and the safety parameter, \(\delta\), is \(SS = \delta \mu\), see Section 2.2. Therefore, the optimal safety stock, \(\delta^*\), can be obtained in the linear system without the discount via Eq. (A.1).

\[
\delta^* = \frac{\sigma}{\mu} \sqrt{1 + Tp} \Phi^{-1} \left[ \frac{b}{b + h} \right]
\]  

(A.1)

Using a similar procedure, we can derive the optimal guaranteed capacity, \(GC^*\) in the linear system. Now we use that the optimal guaranteed capacity for our model of capacity costs is

\[
GC^* = \mu_o + \sigma_o \Phi^{-1} \left[ \frac{p}{p + n} \right]
\]

see Disney and Lambrecht (2008). Here, \(\mu_o = E(\alpha_i)\) is the mean order and \(\sigma_o = \sqrt{\var(\alpha_i)}\) is the standard deviation of the orders. Again, we can particularise this equation specifically for this work. It can be shown that \(\mu_o = E(\alpha_i) = E(\delta) = \mu\), given that lost sales do not occur in our system. Finally, in the linear system, \(BW = \frac{\var(\alpha_i)}{\var(\delta)} = \frac{\sigma^2}{\sigma^2} = 1\), thus \(\sigma_o = \sigma\). Then, the final expression of the optimal guaranteed capacity becomes Eq. (A.2).

\[
GC^* = \mu + \sigma \Phi^{-1} \left[ \frac{p}{p + n} \right]
\]  

(A.2)

References

Ancarani, A., Di Mauro, C., D’Urso, D., 2016. Measuring overconfidence in inventory management decisions. J. Purchas. Supply Manage. 22 (3), 171–180.
Atasu, A., Guide Jr, V.D.R., Van Wassenhove, L.N., 2010. So what if remanufacturing cannibalizes my new product sales? Cal. Manage. Rev. 52 (2), 56–76.
Axäter, S., 2003. Supply chain operations: Serial and distribution inventory systems. Handbook Operat. Rese. Manage. Sci. 11, 525–559.
Bhattacharya, R., Bandyopadhyay, S., 2003. A review of the causes of bullwhip effect in a supply chain. Int. J. Adv. Manuf. Technol. 54 (9–12), 1245–1261.
Burbidge, J.L., 1984. Automated production control with a simulation capability. IFIP Working Papers, ref. WG5(7).
Butman, J., 2002. A pain in the (supply) chain. Harvard Bus. Rev. 80 (5), 31–39.
Cantor, D.E., Katok, E., 2012. Production smoothing in a serial supply chain: a laboratory investigation. Transport. Res. E: Logist. Transport. Rev. 48 (4), 781–794.
Ciancimino, E., Cannella, S., Brucoleti, M., Framinan, J.M., 2012. On the bullwhip avoidance phase: the synchronised supply chain. Eur. J. Oper. Res. 221 (1), 49–63.
Chatfield, D.C., Pritchard, A.M., 2013. Returns and the bullwhip effect. Transport. Res. E: Logist. Transport. Rev. 49 (1), 159–175.
Chen, R.R., Robinson, L.W., 2012. Optimal multiple-breakpoint quantity discount schedules for customers with heterogeneous demands: all-unit or incremental? IIE Trans. 44 (3), 199–214.
Croson, R., Donohue, K., 2006. Behavioral causes of the bullwhip effect and the observed value of inventory information. Manage. Sci. 52 (3), 323–336.
Dejonckheere, J., Disney, S.M., Lambrecht, M.R., Tovill, D.R., 2003. Measuring and avoiding the bullwhip effect: a control theoretic approach. Eur. J. Oper. Res. 147 (3), 567–590.
Disney, S.M., Gaalman, G., Hosoda, T., 2012. Review of stochastic cost functions for production and inventory control. In: Proceedings of the 17th International Working Seminar of Production Economics, pp. 117–128.
