

A numerical study on the effect of the balance assumption in one-warehouse multi-retailer inventory systems

Mustafa K. Doğru · A. G. de Kok · G. J. van Houtum

Published online: 26 October 2010
© Springer Science+Business Media, LLC 2010

Abstract One-warehouse multi-retailer systems under periodic review have been studied extensively in the literature. The optimal policy has not been characterized yet. It would require solving a multi-dimensional dynamic program, which is hard due to the curse of dimensionality. In order to let the dynamic program decompose, researchers often make the so-called *balance assumption*. All available heuristics for periodic review distribution systems are based on some form of this assumption. For these heuristics, often further approximate steps are applied. We investigate the pure effect of the balance assumption in this paper. The balance assumption is the relaxation of a set of constraints in the original dynamic program and yields a lower bound model, which we solve exactly. This gives us a lower bound for the optimal cost of the original model. An upper bound for the true optimal cost is obtained by simulating the optimal policy for the relaxed problem with a slightly modified allocation rule. This modified policy is referred to as the LB heuristic policy. We use the relative gap between the upper and lower bound as a measure to assess the impact of the balance assumption. Based on extensive testing, we identify when the gap is small, and when *not*. For those instances with small gaps, both the lower bound is tight and the performance of the LB heuristic policy is close to the optimal. We also identify many practically relevant settings under which the balance assumption yields large gaps. For these instances, either the lower bound is poor or the LB heuristic policy is far from optimal, or both. In any case, it implies that more

M. K. Doğru (✉)
Alcatel-Lucent Bell Labs, Blanchardstown Industrial Park, Dublin 15, Ireland
e-mail: mustafa.dogru@alcatel-lucent.com

A. G. de Kok · G. J. van Houtum
Industrial Engineering & Innovation Sciences, Technische Universiteit Eindhoven,
5600 MB Eindhoven, Netherlands
e-mail: a.g.de.kok@tue.nl

G. J. van Houtum
e-mail: g.j.v.houtum@tue.nl

research is needed to develop better lower bounds and/or better heuristics for these instances.

Keywords Multi-echelon · Distribution systems · One-warehouse multi-retailer systems · Balance assumption · Bounds · Relaxation

1 Introduction

This paper considers a two-echelon divergent (distribution, arborescent) inventory system where a central stock point, called the warehouse, supplies N downstream stock points, called retailers. Demands of the customers originate at the retailers. The warehouse orders from an external supplier with ample stock and ships to the retailers as needed. There are fixed lead times between the supplier and the warehouse (l_0), and in between the warehouse and the retailers. Any unfulfilled demand of a customer is backlogged. The system is controlled centrally under periodic review.

The model considered is also relevant in a manufacturing context. Suppose an intermediate product/subassembly is stocked at a central location, and used in the manufacturing/assembly of multiple different end products. Production/inventory decisions in such a system can be determined using the one-warehouse multi-retailer model. Moreover, the model can be used to study hierarchical production planning (see de Kok (1990)) and delayed product differentiation (see Aviv and Federgruen (2001)).

The seminal work on periodic review multi-echelon inventory systems is by Clark and Scarf (1960). The authors developed a dynamic programming formulation for the inventory control of an N -echelon serial system in a finite horizon. By introducing the concept of echelon stock and induced penalty cost, they were able to decompose the resulting multi-dimensional dynamic program (DP) into a series of single-dimensional programs (referred to as the *decomposition property*), and characterize the optimal policy, which is an echelon base stock policy. Clark and Scarf (1960) also tried to apply their decomposition technique to divergent structures, but it is not possible to show the decomposition property due to the *allocation (rationing) problem*. (In a one-warehouse multi-retailer setting, in each period the decision maker decides on the amount of stock to keep at the warehouse and ship to each retailer, which is called the allocation problem.) Thus, the optimization problem requires the solution of a multi-dimensional DP with $N + l_0 + 1$ dimensions, which is unrealistic to calculate for a real life application. The structure of the optimal policy for the one-warehouse multi-retailer model is unknown. However, an assumption, known as the *balance assumption*, leads to the decomposition property and a full characterization of the optimal policy. This key assumption is the relaxation of the physical constraint that the inventory positions of the retailers just after the allocation decision are greater than or equal to the inventory positions prior to this decision. Other interpretations of the balance assumption are (1) allowing negative quantities to be shipped to the retailers, (2) permitting immediate return (with no lead time) of stock at any retailer to the

warehouse at no cost, (3) allowing the lateral transshipment (shipments between the retailers) of stock with the lead time of the receiving retailer at no cost.¹

The balance assumption allows the shipment decisions to be just based on the echelon stock of the warehouse. The individual inventory positions of the retailers become irrelevant since negative shipments are permitted. The inventory positions of the retailers under such a shipment scheme represent an *ideal* state where the retailer inventories are said to be in *balance*. However, in a real setting, there may be departures from the ideal state and this contributes to the *imbalance* of the system.

The balance assumption was introduced by Eppen and Schrage (1981) who coined the phrase *allocation assumption* for what we refer to as the balance assumption. In the multi-echelon inventory literature, any multi-period model that includes the allocation problem makes the balance assumption in some form. On one hand, the studies by Federgruen and Zipkin (1984b, c), Erkip et al. (1990), Chen and Zheng (1994), Bollapragada et al. (1998), Diks and de Kok (1998), Cachon and Fisher (2000), and Özer (2003) have equal frequency of ordering and shipment (i.e., the system orders and ships every period), so the system is assumed to be in balance at the beginning of each period. On the other hand, if the warehouse orders less frequently than it ships to the retailers, e.g., Jönsson and Silver (1987), Jackson (1988), Jackson and Muckstadt (1989), McGavin et al. (1993), van der Heijden (1999), then the balance of the system is assumed at each epoch an order is received by the warehouse. Kumar et al. (1995) incorporate another form of the balance assumption.

To our knowledge, all heuristics developed for periodic review divergent systems that benefit from risk pooling² are based on the balance assumption in some form, e.g., Federgruen and Zipkin (1984a), van der Heijden et al. (1997), Diks and de Kok (1999), Kumar and Jacobson (1998), Axsäter et al. (2002), Cao and Silver (2005), Gallego et al. (2007), Kunnumkal and Topaloglu (2008). The majority of these heuristics apply further approximate steps after making the balance assumption. For example, consistent appropriate share and balanced stock heuristic as discussed in van der Heijden et al. (1997) assume linear allocation functions, which ration the on-hand warehouse stock with respect to fixed retailer specific fractions.

There is an established belief in the literature that the balance assumption is a good approximation. In his comprehensive survey study, Axsäter (2003, p. 544) states:

The balance assumption has been used extensively in the inventory literature and has been shown to produce solutions of very good quality in many different situations, ...

¹ In this setting, lateral transshipment does not only imply the shipment of on-hand stock from one retailer to the other, but also includes shipment of stock from one retailer's pipeline inventory to the other.

² Risk pooling is the reduction in the uncertainty faced by the system by carrying a single inventory during the warehouse lead time rather than individual retailer inventories. Virtual allocation proposed by Graves (1996) is not based on the balance assumption, but this method of rationing does not benefit from risk pooling at all and not advocated by the author for practical implementations. See also Axsäter et al. (2002).

The numerical results of the aforementioned references advocate the use of the balance assumption with an exception of Axsäter et al. (2002). We review the literature on the balance assumption in Sect. 2. Although Axsäter et al. (2002) and Gallego et al. (2007) report results that hint the possibility of large errors due to the balance assumption, there is no clear-cut study in the literature that questions the appropriateness of this assumption. Our paper tries to fill this gap by a numerical study conducted over a wide range of parameters.

Our methodology can be explained as follows. The optimization problem under the balance assumption, which we refer to as the *relaxed model* or the *lower bound model*, is a relaxation of the original problem. Hence, the corresponding optimal cost is a lower bound (*LB*) for the true optimal cost. When the optimal base stock levels of the relaxed problem are coupled with a myopic allocation procedure,³ the resulting policy is a feasible heuristic policy, referred to as the *LB heuristic policy*. An estimate for the average expected cost of the LB heuristic policy can be determined by simulation, which serves as an upper bound (*UB*) for the true optimal cost. The relative gap between the bounds ($\epsilon\% = 100 \frac{UB-LB}{LB}$) is used as a measure to assess the impact of the balance assumption. Since the optimal cost is enveloped by *LB* and *UB*, a *small relative gap* implies that the balance assumption leads to a good approximation for that particular input parameter setting: while *LB* is an accurate proxy for the optimal cost, the LB heuristic policy performs well. Hence, we conclude that a small relative gap justifies the use of balance assumption for that setting. On the other hand, when there is a *moderate* or a *large relative gap*, the exact position of the true optimal cost between *LB* and *UB* becomes an important issue. When $\epsilon\%$ is considerable, either *LB* is an inaccurate proxy or the LB heuristic policy is mediocre, or both. If *LB* is tight, then there is a need for a better heuristic. If the cost of the LB heuristic policy is close to the true optimal cost, then this should be established either by developing a better lower bound or determining the optimal cost. As long as it is not demonstrated, neither the LB heuristic policy, nor the heuristics based on the balance assumption can be trusted for the instances with moderate or large gaps.

The following input parameters are considered in the numerical study: holding and penalty costs, warehouse and retailer lead times, the number of retailers, and the mean and the coefficient of variation of the demand processes. We generated two test beds consisting of 2,000 and 3,888 problem instances for identical (in terms of costs, lead times, and demand distributions) and nonidentical retailers cases, respectively. The number of retailers is restricted to two in the nonidentical retailers case.

The results directed us to a clear and complete overview on when the relative gap is *small* and when *not*. In the identical retailers case, the relative gap is small when one of the following conditions holds:

- the coefficient of variation is low or moderate,
- the added value at the warehouse is negligible compared to the ones at the retailers,
- the warehouse lead time is short and the retailer lead times are long.

³ The warehouse on-hand stock is rationed such that the sum of the expected holding and penalty costs of the retailers in the periods the allocated quantities reach their destinations is minimized.

As long as any one condition is satisfied, independent of the extent of the other input parameters, the relative gap is found to be small. In other problem instances, large relative gaps up to 38.6% are observed. The main drivers of large relative gaps are high coefficients of variation and a long warehouse lead time. In the nonidentical retailers case, the relative gap is small when the warehouse lead time is short and lead time of the big retailer (in terms of mean demand) is long. The main determinants of large gaps (up to 186.9%) are positive added value at one retailer and zero at the other, and long warehouse lead time.

Our main contribution to the literature is as follows. To the best of our knowledge, this is the most comprehensive numerical study *specifically* targeting to assess the impact of the balance assumption on the expected long-run cost. We explicitly identify the parameter settings under which the relative gaps are small. We show that the balance assumption may lead to large gaps in many settings, which contradicts with the conviction that it leads to an accurate approximation in almost all cases. The settings with considerable gaps indicate that either *LB* is an inaccurate approximation or the *LB* heuristic policy performs poorly, or both. Further, the available heuristics in the literature, which are based on the balance assumption, cannot be relied on for these problem instances. Our results point out that the scenarios with moderate or large relative gaps are practically relevant, so there is a good prospect for future research on these settings (e.g., to determine the position of the true optimal cost between the bounds). The analysis of the relations between the relative gap and the input parameters led us to new and interesting insights, e.g., shrinkage in the relative gap when retailer lead times increase, the so-called *forwarding-to-the-small-retailer* phenomenon, and the combined effect of retailer size and other input parameters on the relative gap.

The rest of the paper is organized as follows. The literature review on the balance assumption is given in Sect. 2. We present the notation, the model and a brief analysis in Sect. 3. Section 4 is dedicated to the presentation and discussion of the numerical results. A detailed summary of the results and the insights learned are given in Sect. 4.3. We close with a brief conclusion and directions for further research in Sect. 5.

2 Literature on the balance assumption

As discussed in Sect. 1, the balance assumption is utilized extensively in the literature, but there are only a few studies that analyze the quality of this assumption. Zipkin (1984) proposed the first analytical approach, which uses dynamic programming. In a relatively simple setting (zero lead times for orders and shipments, warehouse as a cross-docking point), the author developed a DP that accounts for the imbalance in the system. In the spirit of the induced penalty cost approach of Clark and Scarf (1960), the system cost is modelled consisting of two components: cost obtained under the balance assumption, and an additional cost that is a consequence of the imbalance in the system. A numerical study shows the accuracy of the approximation for the limited number of scenarios

considered. It is concluded that imbalance can be considerable when demand variances are large. The studies by van Donselaar and Wijngaard (1987), and van Donselaar (1990) investigate the effect of imbalance on P_1 service level, i.e., probability of stockout. In the former study, the results of a numerical study conclude the little impact of imbalance on the system service level. We believe that the numerical study conducted is rather restricted to come to such a conclusion. The latter paper, incorporates the effect of batch sizes on the imbalance of retailer inventories.

Federgruen and Zipkin (1984a), Kumar and Jacobson (1998), Axsäter et al. (2002), and Gallego et al. (2007) developed heuristics for the control of one-warehouse multi-retailer systems. Since the optimal policy and the associated cost is unknown, instead of comparing the cost of their heuristics to the true optimal cost, they all used the relative gap between the optimal cost of the system under the balance assumption and the cost of the heuristics (found by simulation) as a performance measure. The numerical results of these papers give rough ideas on the effect of the balance assumption on the system-wide cost. The results by Federgruen and Zipkin, and Kumar and Jacobson indicate that the gaps are small in general, with maximum relative gaps of 6.42 and 0.78%, respectively. Axsäter et al. considered a total of 68 scenarios where several of them exhibit high relative gaps with a maximum of 226.5%. Similarly, Gallego et al. report cases with high relative gaps, especially when some of the retailers have negligible added value.

In this study, we follow the same line of thought of Federgruen and Zipkin (FZ), Kumar and Jacobson (KJ), Axsäter et al. (AMS) and Gallego et al. (GOZ), but our study is different in certain aspects. The main dissimilarities are:

- (1) The focal point of these four studies is the performance of their heuristics; we focus particularly on the effect of the balance assumption. Hence, while we consider 5888 scenarios, FZ, KJ, AMS and GOZ have 138, 68, 68 and 409 scenarios, respectively.
- (2) FZ consider a fixed cost for ordering from the external supplier and AMS assume a fixed batch size for warehouse orders; we do not consider these issues.
- (3) FZ and KJ consider a warehouse that cannot hold stock; we relax this restriction.
- (4) FZ, and KJ assume normal demands at the retailers; when the coefficient of variation is higher than 0.5, the normal distribution assumption creates complications due to significant probability of negative demand appearing in the analysis. In real life, the coefficient of variation is higher than 0.5 in many cases, and can even be more than three, see Dođru (2006) for real life examples. GOZ consider Poisson demands. The Poisson distribution has a coefficient of variation that is equal to the inverse of the square root of the mean. Hence, demand streams with high coefficient of variation can only be created with very low arrival rates. We model one-period demands as mixtures of Erlang distributions, which allows us to incorporate instances with

coefficients of variation up to three, and change the coefficient of variation while keeping the mean demand constant.

- (5) For nonidentical retailers, FZ require the holding and penalty costs to be proportional, and KJ assume identical cost parameters among the retailers. We relax such restrictions.
- (6) GOZ consider a continuous-review model while the rest, including ours, study a periodic-review setting.

Especially, the points (4) and (5) allow us to consider some extreme scenarios (e.g., coefficient of variation of demand equal to 3, or added value of zero at a retailer) for the purpose of finding out the extent of the impact of the balance assumption.

3 Model and analysis

Consider a one-warehouse multi-retailer inventory system controlled centrally under a periodic review setting. The warehouse (indexed as stock point 0) places *orders* to an exogenous supplier with ample stock, and the retailers are replenished by *shipments* from the warehouse. The warehouse is allowed to hold stock. All lead times are assumed to be fixed. The retailers face the stochastic i.i.d. demand of the customers that are stationary and continuous on $(0, \infty)$ with no probability mass at zero. Costs consist of linear inventory holding and penalty costs. Time is divided into periods of equal length and we assume that the following sequence of events takes place during a period:

- inventory levels are observed, the warehouse replenishes its inventory from the exogenous supplier, and any order due to arrive is received (at the beginning of the period);
- the on-hand inventory at the warehouse is allocated and the shipments due to arrive are received by the retailers (at the beginning of the period);
- demand occurs;
- holding and penalty costs are assessed on the period ending inventory and backorder levels (at the end of the period).

The objective is to minimize the expected holding and backordering costs of the system in the long-run. We use the conventional definition of *echelon stock* of a stock point, which is the stock at that stock point plus in transit to or on-hand at any stock point downstream minus the backorders. *Echelon inventory position* of a stock point is the echelon stock plus all inventory that is in-transit to this stock point. The notation used is given in Table 1.

3.1 Dynamics of the system

Before discussing the dynamics of the system, we introduce the concept of *cost attached to an echelon*. The total cost of the system at the end of an arbitrary period t is

Table 1 The notation for the study

\mathbb{R}	= set of real numbers
\mathbb{Z}	= set of integer numbers; $\mathbb{Z}^+ = \{1, 2, \dots\}$ and $\mathbb{Z}_0^+ = \mathbb{Z}^+ \cup \{0\}$
t	= index for time. Period t is defined as the time interval between epochs t and $t + 1$ for $t \in \mathbb{Z}_0^+$
N	= number of retailers, $N \in \mathbb{Z}^+$
i	= index for stock points, $i = 0$ is the warehouse and $i = 1, 2, \dots, N$ are the retailers
J	= set of retailers, $J = \{1, 2, \dots, N\}$
h_i	= additional inventory holding cost parameter for stock point i . At the end of a period: (i) cost $h_0 \geq 0$ is charged for each unit on stock at the warehouse or in transit to any of the retailers, (ii) cost $h_0 + h_i$ is charged for each unit on stock at retailer i ($h_i \geq 0, i \in J$).
p_i	= penalty cost parameter for retailer i . A cost p_i is charged for each unit of backlog at the end of a period at retailer i ($p_i > 0, i \in J$)
l_i	= lead time parameter for stock point i ($l_i \in \mathbb{Z}_0^+$ for $i \in J$, and $l_0 \in \mathbb{Z}^+$)
μ_i	= mean one-period demand faced by retailer i ($\mu_i > 0, i \in J$)
σ_i	= standard deviation of one-period demand faced by retailer i
μ_0	= mean one-period demand faced by the system, i.e., $\mu_0 = \sum_{i \in J} \mu_i$
cv_i	= coefficient of variation of one-period demand faced by retailer i , $cv_i = \sigma_i/\mu_i$
$F_i^{(l)}$	= cumulative distribution function of l -period demand of retailer i
$D_i(t, t + s)$	= random demand faced by retailer i during the periods $t, t + 1, \dots, t + s$ for $t, s \in \mathbb{Z}_0^+$
$D_0(t, t + s)$	= aggregate random demand faced by the system during the periods $t, t + 1, \dots, t + s$, i.e., $D_0(t, t + s) = \sum_{i \in J} D_i(t, t + s)$, for $t, s \in \mathbb{Z}_0^+$
$I_i(t)$	= echelon stock of stock point i at the beginning of period t just after the receipt of the incoming order (if $i = 0$) or shipment ($i \in J$)
$\hat{I}_i(t)$	= echelon stock of stock point i at the end of period t
$IP_i(t)$	= echelon inventory position of stock point i at the beginning of period t just after the ordering (if $i = 0$) or shipment ($i \in J$) decision
$\hat{IP}_i(t)$	= echelon inventory position of retailer i at the end of period t = echelon inventory position of retailer i at the beginning of period $t + 1$ just before the shipment decision

$$h_0 \left(\hat{I}_0(t) - \sum_{i \in J} \hat{I}_i(t) \right) + \sum_{i \in J} (h_0 + h_i) \hat{I}_i^+(t) + \sum_{i \in J} p_i \hat{I}_i^-(t),$$

where $a^+ = \max\{0, a\}$ and $a^- = -\min\{0, a\}$ for $a \in \mathbb{R}$. Substituting $\hat{I}_i(t) = \hat{I}_i^+(t) - \hat{I}_i^-(t)$, rearranging the terms, and then using the identity $\hat{I}_i^+(t) = \hat{I}_i(t) + \hat{I}_i^-(t)$ yields

$$\begin{aligned} & h_0 \left(\hat{I}_0(t) - \sum_{i \in J} \hat{I}_i(t) \right) + \sum_{i \in J} (h_0 + h_i) \hat{I}_i^+(t) + \sum_{i \in J} p_i \hat{I}_i^-(t) \\ &= h_0 \hat{I}_0(t) + \sum_{i \in J} h_i \hat{I}_i^+(t) + \sum_{i \in J} (h_0 + p_i) \hat{I}_i^-(t) \\ &= h_0 \hat{I}_0(t) + \sum_{i \in J} h_i \hat{I}_i(t) + \sum_{i \in J} (h_0 + h_i + p_i) \hat{I}_i^-(t). \end{aligned}$$

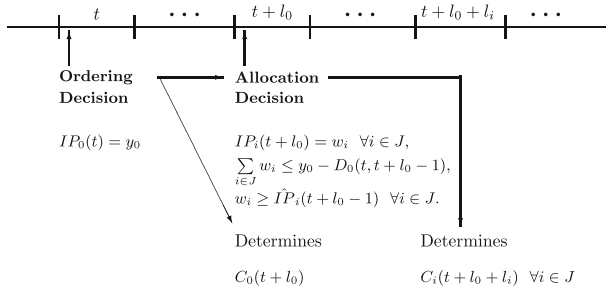


Fig. 1 The consequences of an order placed in period t

We define $C_0(t) = h_0 \hat{I}_0(t)$ as the cost attached to the echelon of the warehouse (echelon 0) at the end of period t , and $C_i(t) = h_i \hat{I}_i(t) + (h_0 + h_i + p_i) \hat{I}_i^-(t)$ as the cost attached to the echelon of retailer $i \in J$ at the end of period t .

Consider the following two decisions and their effects on the expected costs, which starts with an order given to the supplier in period $t, t \in \mathbb{Z}_0^+$. Figure 1 illustrates the arguments.

- Ordering Decision:** At the beginning of period t , the warehouse gives an order that raises the inventory position of the system up to y_0 , i.e., $IP_0(t) = y_0$. (Note that y_0 is not bounded from above due to the assumption that the supplier has ample stock). The order materializes at the beginning of period $t + l_0$ and the available stock in the echelon of the warehouse at that epoch is $y_0 - D_0(t, t + l_0 - 1)$. First, this ordering decision directly affects the cost attached to the echelon of the warehouse at the end of period $t + l_0$. The expectation of $C_0(t + l_0)$ conditioned on $IP_0(t) = y_0$

$$E[C_0(t + l_0) | IP_0(t) = y_0] = E[h_0(y_0 - D_0(t, t + l_0))] = h_0(y_0 - (l_0 + 1)\mu_0).$$

Second, it puts an upper bound on the level to which one can increase the aggregate echelon inventory position of the retailers in period $t + l_0$, i.e., $\sum_{i \in J} IP_i(t + l_0) \leq y_0 - D_0(t, t + l_0 - 1)$.

- Allocation Decision:** At the beginning of period $t + l_0$, shipments to the retailers are determined; equivalently, the system-wide stock is allocated among all stock points. That is to say, the decisions of how much to send to each retailer and how much to keep at the warehouse are made. The inventory position of retailer i is increased up to w_i , i.e., $IP_i(t + l_0) = w_i$. Note that retailer replenishment decisions should be made subject to $\hat{IP}_i(t + l_0 - 1) \leq w_i$ for all $i \in J$, and $\sum_{i \in J} w_i \leq y_0 - D_0(t, t + l_0 - 1)$. These decisions directly affect the cost attached to echelon i at the end of period $t + l_0 + l_i$; let $G_i(w_i)$ be the expected value of this cost conditioned on $IP_i(t + l_0) = w_i$:

$$\begin{aligned}
 G_i(w_i) &= E[C_i(t + l_0 + l_i) | IP_i(t + l_0) = w_i] \\
 &= E[h_i(w_i - D_i(t + l_0, t + l_0 + l_i)) + (h_0 + h_i + p_i)(w_i - D_i(t + l_0, t + l_0 + l_i))^-] \\
 &= h_i(w_i - (l_i + 1)\mu_i) + (h_0 + h_i + p_i)E[(w_i - D_i(t + l_0, t + l_0 + l_i))^-],
 \end{aligned}$$

for all $i \in J$. Define $C_{cyc}(t)$ as the sum of the costs that are a consequence of the warehouse’s ordering decision in period t :

$$C_{cyc}(t) = C_0(t + l_0) + \sum_{i \in J} C_i(t + l_0 + l_i). \tag{1}$$

We refer to $C_{cyc}(t)$ as the *cycle cost of period t* .

Let Π denote the set of all ordering/allocation policies. Define $g_\pi(\mathbf{x}_0)$ as the long-run average expected cost of policy π given an initial state \mathbf{x}_0 , which is an $N + l_0 + 1$ dimensional vector showing echelon stock of the warehouse, the echelon inventory positions of the retailers, and all the orders in-transit at time zero just before the decisions are made. The expected long-run average cost of any policy $\pi \in \Pi$ is simply the average of the expected value of the sum of all costs attached to each echelon in an infinite horizon:

$$g_\pi(\mathbf{x}_0) \stackrel{\text{def}}{=} \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{T-1} \sum_{i=0}^N C_i(t) \right].$$

Assuming that the limit exists and finite⁴, the expression above can be streamlined as

$$\begin{aligned} g_\pi(\mathbf{x}_0) &= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{T-1} \sum_{i=0}^N C_i(t) \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=0}^{l_0-1} C_0(t) + \sum_{i \in J} \sum_{t=0}^{l_0+l_i-1} C_i(t) \right. \\ &\quad \left. + \sum_{t=0}^{T-1} C_{cyc}(t) - \sum_{t=T}^{T+l_0-1} C_0(t) - \sum_{i \in J} \sum_{t=T}^{T+l_0+l_i-1} C_i(t) \right] \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[C_{cyc}(t)], \end{aligned}$$

where substituting (1) yields the second equality, and the third equality follows from all terms except the third term in the second equality being finite.

The optimization problem for a given initial state \mathbf{x}_0 is

$$\min_{\pi \in \Pi} g_\pi(\mathbf{x}_0), \tag{2}$$

and we denote the optimal cost by g^* . The minimization problem given above is intricate since the decisions are highly interdependent. In the next subsection, we introduce the myopic allocation problem and discuss the balance assumption.

We close this subsection with a final remark on variable ordering and transportation costs. Let c_0 be the variable ordering cost per unit at the warehouse

⁴ The existence and finiteness of the limit may not hold for any given policy, especially for the ones that do not order enough to satisfy demand, but any stationary policy that generates a unichain Markov chain meets this requirement. We are interested in such policies.

and c_i be the unit transportation cost between the warehouse and retailer i . In an infinite horizon, μ_0 is the average quantity ordered in a period and μ_i is the average quantity shipped to retailer i per period. The average variable cost associated to ordering and shipments is $c_0\mu_0 + \sum_{i \in J} c_i \mu_i$ and constant in an infinite horizon average cost model, hence, it does not influence the ordering and allocation decisions. That is why it is common practice in the literature to omit these costs in the analysis.

3.2 Myopic allocation problem

Consider the sequence of decisions and the resulting costs as a result of increasing the inventory position of the system up to y_0 at the beginning of period t , $t \in \mathbb{Z}_0^+$. Assume that the stock in the system at the beginning of period $t + l_0$ (i.e., $y_0 - D_0(t, t + l_0 - 1)$) is allocated among all stock points such that the sum of the expected holding and backordering costs of the retailers in the periods during which the allocated quantities reach their destinations (i.e. period $t + l_0 + l_i$ for retailer i) is minimized. Since the effect of the allocation decision on the subsequent periods is not considered, this way of rationing is called *myopic allocation*. The mathematical formulation of the myopic allocation problem of period $t + l_0$ is as follows:

$$\min_{w_i, \forall i \in J} \sum_{i \in J} G_i(w_i) \tag{3}$$

$$s.t. \sum_{i \in J} w_i \leq y_0 - D_0(t, t + l_0 - 1) \tag{4}$$

$$\hat{I}P_i(t + l_0 - 1) \leq w_i \quad \forall i \in J, \tag{5}$$

where both constraints serve for the physical balance of the stocks. While (5) restricts the allocated quantities to be nonnegative, (4) assures that the sum of the allocated quantities cannot exceed the stock in the system. The solution of the myopic allocation problem depends on the amount to allocate, $y_0 - D_0(t, t + l_0 - 1)$, and each retailer’s inventory position, $\hat{I}P_i(t + l_0 - 1)$.

Although myopic allocation allows the allocation decisions to be made independent of the future allocation and ordering decisions, it still depends on previous periods’ decisions due to (5). Consider a relaxed version of the myopic allocation problem where (5) is omitted. This is equivalent to assuming that the quantities allocated to the retailers may be negative. We refer to this assumption as the *balance assumption*. This relaxation helps the analysis as follows. In the absence of (5), $C_{cyc}(t)$ depends *only* on the ordering and allocation decisions that start with the order placed in period t ; not on decisions of other periods. Then, it can be proved that a myopic allocation is optimal for (2), and the decomposition technique of Clark and Scarf (1960) can be applied to characterize the optimal ordering and allocation decisions (see Federgruen and Zipkin (1984b); Federgruen and Zipkin (1984c)). In addition, the myopic allocation problem becomes a convex and separable nonlinear optimization problem subject to a linear constraint, which can be solved by the Lagrangian technique.

Let $z_i : \mathbb{R} \rightarrow \mathbb{R}, i \in J$ be an allocation function such that $z_i(a)$ is the portion of a allocated to retailer i for $a \in \mathbb{R}$. Assume that the system-wide stock at time $t \in \mathbb{Z}_0^+$ is $x \in \mathbb{R}$, i.e., $I_0(t) = x$. The myopic allocation problem of period t under the balance assumption may be rewritten as:

$$\min_{z_i(x), \forall i \in J} \left\{ \sum_{i \in J} G_i(z_i(x)) : \sum_{i \in J} z_i(x) \leq x, z_i(x) \in \mathbb{R} \forall i \in J \right\} \tag{6}$$

where a solution is denoted by $\{z_i(x)\}_{i \in J}$. Let $\{z_i^*(x)\}_{i \in J}$ be an optimal solution of (6) for a given x . Define

- \mathbf{z} = set of allocation functions, i.e., $\{z_i(x), x \in \mathbb{R}\}_{i \in J}$.
- \mathbf{z}^* = set of optimal allocation functions, i.e., $\{z_i^*(x), x \in \mathbb{R}\}_{i \in J}$.

Note that $G_i(\cdot)$ is a newsvendor function, thus, is convex. Let $y_i^* = \arg \min G_i(y_i)$, and verify that y_i^* satisfies the following well-known newsvendor equation:

$$F_i^{(l_i+1)}(y_i^*) = \frac{h_0 + p_i}{h_0 + h_i + p_i} \quad \text{for } i = 1, 2, \dots, N. \tag{7}$$

The interpretation of the expression above is that y_i^* corresponds to a level at which the probability of being out of stock is $\frac{h_i}{h_0 + h_i + p_i}$ for retailer i .

The objective function in (6) is separable and consists of N convex components. The solution $\{z_i^*(x) = y_i^*\}_{i \in J}$ is optimal for (6) when the constraint is not binding, i.e., for $x > \sum_{i \in J} y_i^*$. For $x < \sum_{i \in J} y_i^*$, Lagrangian relaxation can be used to find an optimal solution for (6).

3.3 Lower bound model

The balance assumption relaxes a constraint in the original optimization problem. Thus, (6) is used for allocation instead of (3)–(5). Due to this relaxation, the expected average cost obtained by an optimal policy derived under the balance assumption serves as a lower bound for (2).

Denote a *base stock policy* by a tuple (y_0, \mathbf{z}) , where y_0 is the target echelon inventory position of the warehouse, and $z_i(x)$ for all $i \in J$ are the target levels of the retailers when the system-wide on-hand stock is x . The decisions are made such that, at the beginning of each period $t \in \mathbb{Z}_0^+$:

- the echelon inventory position of the warehouse is increased to y_0 , i.e., $IP_0(t) = y_0$,
- the inventory position of retailer i is raised to $z_i(I_0(t))$ where $\{z_i(I_0(t))\}_{i \in J}$ is a solution that satisfies the constraints of (6).

Consider the expected cycle cost of period t . As mentioned before, under the balance assumption, $E[C_{cyc}(t)]$ depends only on the ordering and allocation decisions that start with an order placed in period t . As a result, it can be shown that a base stock policy (y_0^*, \mathbf{z}^*) is the optimal replenishment policy with y_0^* satisfying the following newsvendor-type equations:

$$\int_0^{\infty} F_i^{(l_i+1)}(z_i^*(y_0^* - u)) dF_0^{(l_0)}(u) = \frac{P_i}{h_0 + h_i + p_i} \quad \forall i \in J, \quad (8)$$

(cf. Diks and de Kok (1998)). Equation (8) indicates that y_0^* is the level at which the probability of being out of stock at retailer i is $\frac{h_0+h_i}{h_0+h_i+p_i}$ assuming that the system never experiences imbalance.

Due to the fact that the warehouse order-up-to level (y_0^*) and optimal allocation functions (\mathbf{z}^*) are independent of time, policy (y_0^*, \mathbf{z}^*) can be applied to optimize each period's cycle cost within the horizon. Thus, the expected average cost obtained by following (y_0^*, \mathbf{z}^*) is a lower bound (*LB*) for (2).

3.4 Constructing an upper bound

Consider base stock policy (y_0^*, \mathbf{z}^*), which is discussed in Sect. 3.3. When policy (y_0^*, \mathbf{z}^*) is implemented, there might be negative allocations, which make it infeasible to apply the solution of (6). Consider the following feasible policy, which we refer to as *LB heuristic policy* since it is based on the lower bound model:

- **ordering:** the inventory position of the warehouse is raised to y_0^* at the beginning of each period,
- **allocation:** the myopic allocation problem given in (3)–(5) is solved to determine the shipment quantities.

The expected cost of the *LB heuristic policy* is an upper bound (*UB*) for g^* . Unfortunately, it is not that straightforward to calculate the expected average cost of the *LB heuristic policy*, so an estimate can be obtained by simulating the system under this policy. (See also Axsäter et al. (2002, p. 79).)

3.5 The gap between *LB* and *UB*

For a given problem instance, the optimal inventory control parameters can be calculated for the lower bound model since analytical results are available. An estimate for an upper bound can be determined by simulating the *LB heuristic policy*. Since g^* lies between *LB* and *UB*, the gap can be used to measure the impact of the balance assumption on the expected system-wide cost. Recall that $\epsilon\% = 100 \frac{UB-LB}{LB}$ where $\epsilon\%$ is defined as the relative gap. If $\epsilon\%$ is small for a problem instance, then one can conclude that the balance assumption is not restrictive for that setting. While *LB* is an accurate proxy for g^* , the *LB heuristic policy* is a good heuristic. On the other hand, a moderate or a large $\epsilon\%$ requires further investigation, which is not the purpose of this paper (we have already discussed this issue in Sect. 1; see also Sect. 5).

4 Numerical study and results

Our main objective is to identify the input parameter settings for which the resulting relative gaps are small and for which they are not. In addition, we try to illustrate the effect of lead times, holding and penalty costs, and the demand processes of the retailers on the relative gap. In order to achieve these goals, two test beds are generated; one for the case where retailers are identical, and the other for nonidentical retailers.

We use the words *scenario* and *problem instance* interchangeably for each combination of the input parameters N, l_0, h_0 and $h_i, p_i, \mu_i, l_i, cv_i$ for $i = 1, \dots, N$. For the calculation of UB , each scenario was simulated for at least 200 replications and terminated as soon as the width of a 95% confidence interval was within 1% of the estimated average cost.

We assume that the underlying demand distribution of a retailer in a given period is a mixture of Erlang distributions with the same scale parameter. An Erlang- k distributed random variable (which is a sum of $k \in \mathbb{Z}^+$ independent exponentially distributed random variables with the same mean λ^{-1}) is denoted by $E_{k,\lambda}$ and the cumulative probability distribution function is $E_{k,\lambda}(x) = 1 - \sum_{j=0}^{k-1} \frac{(\lambda x)^j}{j!} e^{-\lambda x}$ for $x \geq 0$.

When $cv_i \leq 1$, a mixture of $E_{k-1,\lambda}$ and $E_{k,\lambda}$ distributions is fitted for the one-period demand distribution of retailer i ; i.e., $F_i^{(1)}(x) = pE_{k-1,\lambda}(x) + (1 - p)E_{k,\lambda}(x)$ for $x \geq 0$. The parameters $k \geq 2, 0 \leq p \leq 1$ and λ are chosen such that:

$$\frac{1}{k} \leq cv_i^2 \leq \frac{1}{k-1}, \quad p = \frac{1}{1 + cv_i^2} \left[kcv_i^2 - \sqrt{k(1 + cv_i^2) - k^2cv_i^2} \right], \quad \lambda = \frac{k - p}{\mu_i}.$$

When $cv_i > 1$, a mixture of $E_{1,\lambda}$ and $E_{k,\lambda}$ distributions is fitted; i.e., $F_i^{(1)}(x) = pE_{1,\lambda}(x) + (1 - p)E_{k,\lambda}(x)$ for $x \geq 0$. The smallest $k \geq 3$ that satisfies $cv_i^2 \leq \frac{k^2+4}{4k}$ is chosen. The values of p and λ are determined by:

$$p = \frac{2kcv_i^2 + k - 2 - \sqrt{k^2 + 4 - 4kcv_i^2}}{2(k - 1)(1 + cv_i^2)}, \quad \lambda = \frac{p + k(1 - p)}{\mu_i}.$$

(See Tijms (2003, pp. 444–446) for details.)

4.1 Identical retailers

The retailers are identical in terms of lead time, holding and penalty costs, and the demand processes, i.e., $l_i = l_{i+1}, h_i = h_{i+1}, p_i = p_{i+1}, \mu_i = \mu_{i+1}$, and $cv_i = cv_{i+1}$ for $i = 1, \dots, N - 1$. Without loss of generality, the mean demand (μ_i) and the holding cost ($h_0 + h_i$) at each retailer i is kept at 1 for all scenarios. The following set of parameters is used (for $i = 1, 2, \dots, N$):

$$\begin{aligned} h_i &= 0, 0.1, 0.5, 0.9, 0.99 & p_i &= 4, 9, 19, 99 & N &= 2, 3, 4, 5 \\ (l_0, l_i) &= (1, 1), (1, 3), (3, 1), (1, 5), (5, 1) & cv_i &= 0.25, 0.5, 1, 2, 3. \end{aligned}$$

The penalty costs are chosen to assure no-stockout probabilities of 80, 90, 95, and 99% at each retailer under an optimal policy in the lower bound model. Additional inventory holding cost at a stock point reflects the added value at that point. Notice that as h_i increases from 0 to 0.99, the added value at the warehouse decreases since $h_0 + h_i = 1$. A full factorial design is used to generate a test bed that consists of 2,000 problem instances.

During the simulation runs for computing UB , the probability of imbalance was also estimated. The probability of imbalance is defined as the fraction of periods in which a negative quantity is allocated to a retailer when (6) is solved.

We use the analysis of variance to test the impact of the factors ($cv_i, h_i, p_i, N, l_0, l_i$) and their two-factor interactions on $\epsilon\%$. The results of the analysis showed that there is no statistical evidence suggesting N and its two-factor interactions with the other factors have affect on $\epsilon\%$. Also, the interaction between some other factors were found to be not statistically significant. Table 2 summarizes the results of the analysis of variance with statistically significant factors and interactions. All effects considered in the model are significant at 99% level. The R^2 and adjusted R^2 for the model are 0.76 and 0.75, respectively. The ANOVA table suggests that most of the variance in the data set can be explained by the main effects of cv_i, h_i and l_0 , and their interactions. For example, the largest partial sum of squares in the model is due to the interaction between cv_i and h_i , which implies that the effect of the coefficient of variation on the relative gap varies as a function of the additional holding cost at the retailers. Thus, we give the statistics related to the data in two-by-two tables, each corresponding to one interaction in the ANOVA model. The results are given in Tables 3 and 4. For example, the first part of the table shows the effect of the coefficient of variation of demand and the additional holding cost at the retailers. The upper left corner of the table gives various statistics for a set of 80 problem instances having $cv_i = 0.25$ and $h_i = 0$ for $i = 1, 2, \dots, N$. The tabulated *statistics* are average, standard deviation and maximum relative gap (denoted by *ave. $\epsilon\%$* , *std. $\epsilon\%$* and *max. $\epsilon\%$* , respectively), and average, standard deviation and maximum probability of imbalance (denoted by *ave. π* , *std. π* and *max π* , respectively).

The main findings are summarized below.

1. There is statistical evidence that all parameters considered in the numerical experiments except the number of retailers (cv_i, h_i, p_i, l_0, l_i) have impact on the relative gap. Further, some two-way interactions of these parameters are also found to have an effect on $\epsilon\%$ as shown in Table 2.
2. In order to identify the impact of the balance assumption, we looked at *max $\epsilon\%$* and *ave. $\epsilon\%$* values in Tables 3 and 4. An input parameter resulting in *max $\epsilon\%$* ≤ 5 and *ave. $\epsilon\%$* ≤ 1 is considered to be a setting with a small relative gap. The results for identical retailers show that the relative gap is small when
 - the coefficient of variation of the retailers is low or moderate (i.e., $cv_i = 0.25, 0.5$), or
 - the added value at the retailers is very high compared to the one at the warehouse (i.e., $h_i = 0.99$).

Table 2 ANOVA table for fixed-effects model for identical retailers case

Source	Partial SS	MS	F_0	$Prob > F_0$
cv_i	8331.36	2082.84	255.28	0.00
h_i	2782.04	695.51	85.24	0.00
p_i	317.25	105.75	12.96	0.00
l_0	1023.81	511.90	62.74	0.00
l_i	493.90	246.95	30.27	0.00
$cv_i * h_i$	8506.10	531.63	65.16	0.00
$cv_i * p_i$	1736.26	144.69	17.73	0.00
$cv_i * l_0$	2502.67	312.83	38.34	0.00
$cv_i * l_i$	484.70	60.59	7.43	0.00
$h_i * p_i$	536.78	44.73	5.48	0.00
$h_i * l_0$	945.81	118.23	14.49	0.00
$h_i * l_i$	284.72	35.59	4.36	0.00
Residual	15600.39	8.16		
Total	65100.88	32.57		

Especially in scenarios with low coefficient of variation, the relative gap is negligible, see $max \epsilon\%$ and $ave. \epsilon\%$ figures for $cv_i = 0.25$ and 0.5 . There are also some restricted regions where the relative gap is considered to be small, e.g., coefficient of variation of 1 and short warehouse lead time ($l_0 = 1$), but we leave the discussions for later.

Out of the 2,000 problem instances studied, 1,713, 1,487 and 1,275 of them have $\epsilon\%$ less than or equal to 5, 2, and 1, respectively. When the $\epsilon\%$ values (of all instances) are sorted, the scenarios with high relative gap figures have the following common parameter settings:

- 40 scenarios with $\epsilon\% > 25$; all having $cv_i = 3$, $l_i = 1$, $l_0 = 3$ or 5 , and $h_i = 0, 0.1$ or 0.5 ,
- 12 scenarios with $25 \geq \epsilon\% > 20$; all having $cv_i = 3$, $l_i = 1$, and $h_i = 0, 0.1$ or 0.5 ,
- 36 scenarios with $20 \geq \epsilon\% > 15$; all having $cv_i = 2$ or 3 , $l_i = 1$, and $h_i = 0, 0.1$ or 0.5 ,
- 70 scenarios with $15 \geq \epsilon\% > 10$; all having $cv_i = 2$ or 3 , and $l_i = 1$.

The common factor in settings above is high coefficient of variation and short retailer lead times. These observations and the figures in Tables 3 and 4 suggest that when a high coefficient of variation is coupled with moderate or low added value at the retailers, a long warehouse lead time and short retailer lead times, the gaps tend to be high.

3. We also analyzed the relationship between the relative gap and the imbalance probability. The Pearson correlation coefficient is estimated as 0.659, which indicates a positive correlation. Further, the imbalance probabilities are plotted against the corresponding $\epsilon\%$ values in Fig. 2, where a different symbol is used for each coefficient of variation value. First, the results show that a high

Table 3 The summary of the results from scenarios with identical retailers for different cv_i levels

cv_i	h_i										p_i					l_0					l_i					Total
	0.00		0.10		0.50		0.90		0.99		4	9	19	99	1	3	5	1	3	5	1	3	5			
0.25	ave. $\epsilon\%$	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	
	std. $\epsilon\%$	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.01	0.02	0.03	0.02	
	max $\epsilon\%$	0.07	0.07	0.09	0.14	0.15	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.15	0.15	0.07	0.06	0.07	0.06	0.07	0.06	0.07	0.13	0.15	0.15	
	ave. π	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	std. π	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
	max π	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.50	ave. $\epsilon\%$	0.06	0.06	0.04	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.07	0.04	0.05	0.03	0.05	0.03	0.05	0.07	0.05	0.03	0.02	0.04	
	std. $\epsilon\%$	0.08	0.08	0.05	0.03	0.03	0.03	0.03	0.03	0.04	0.04	0.04	0.05	0.09	0.06	0.07	0.05	0.07	0.05	0.07	0.05	0.05	0.05	0.05	0.06	
	max $\epsilon\%$	0.35	0.36	0.20	0.19	0.18	0.17	0.23	0.26	0.26	0.17	0.23	0.26	0.36	0.36	0.33	0.27	0.36	0.33	0.27	0.36	0.21	0.19	0.36	0.36	
	ave. π	0.04	0.04	0.03	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	
	std. π	0.00	0.01	0.01	0.00	0.00	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	
	max π	0.05	0.05	0.04	0.01	0.00	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.05	
1	ave. $\epsilon\%$	1.51	1.52	1.27	0.31	0.05	0.75	0.86	0.96	0.71	0.84	0.95	1.18	1.17	0.73	1.33	1.15	1.25	1.15	1.07	1.25	0.57	0.36	0.93		
	std. $\epsilon\%$	1.03	1.01	0.66	0.14	0.06	0.71	0.84	0.95	0.71	0.84	0.95	1.18	1.17	0.77	1.17	1.01	1.07	1.01	1.07	0.46	0.29	0.95	0.95		
	max $\epsilon\%$	4.93	4.91	3.28	0.66	0.27	2.45	3.07	3.66	3.66	2.45	3.07	3.66	4.93	3.64	4.93	4.50	4.93	4.50	4.93	1.83	1.29	4.93	4.93		
	ave. π	0.34	0.34	0.28	0.08	0.01	0.22	0.21	0.21	0.21	0.22	0.21	0.21	0.20	0.22	0.22	0.20	0.19	0.20	0.19	0.20	0.22	0.23	0.21		
	std. π	0.05	0.05	0.09	0.05	0.00	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.16	0.16	0.15	0.15	0.14	0.15		
	max π	0.41	0.41	0.38	0.22	0.02	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.41	0.38	0.41	0.41	0.41	0.41	0.41	0.41	0.38	0.38	0.41		
2	ave. $\epsilon\%$	6.86	6.94	5.84	1.33	0.09	5.10	3.86	3.34	3.34	3.86	3.34	4.55	4.55	2.63	6.61	6.57	5.92	6.57	5.92	1.95	1.36	4.21	4.21		
	std. $\epsilon\%$	4.64	4.66	3.62	1.08	0.09	5.45	3.55	3.35	3.35	3.55	3.35	4.99	4.99	2.77	5.43	5.35	5.01	5.35	5.01	1.31	0.85	4.46	4.46		
	max $\epsilon\%$	16.53	16.63	15.28	5.39	0.33	15.28	11.66	12.79	12.79	11.66	12.79	16.63	16.63	12.37	15.44	16.63	16.63	15.44	16.63	3.92	2.65	16.63	16.63		
	ave. π	0.48	0.48	0.42	0.16	0.02	0.32	0.33	0.30	0.30	0.32	0.33	0.29	0.29	0.29	0.34	0.33	0.32	0.33	0.32	0.30	0.29	0.29	0.31		
	std. π	0.10	0.10	0.10	0.10	0.02	0.22	0.20	0.20	0.20	0.22	0.20	0.21	0.21	0.17	0.24	0.26	0.23	0.26	0.23	0.17	0.17	0.17	0.21		
	max π	0.64	0.64	0.63	0.47	0.09	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.49	0.62	0.64	0.64	0.64	0.64	0.49	0.49	0.49	0.64		

Table 3 continued

CV_j	h_i	P_i					l_0					l_i					Total
		0.00	0.10	0.50	0.90	0.99	4	9	19	99	1	3	5	1	3	5	
3	ave. $\epsilon\%$	12.82	12.92	9.43	2.41	0.24	6.45	11.30	6.37	6.14	3.68	12.64	14.15	10.81	3.04	2.36	7.56
	std. $\epsilon\%$	11.12	11.19	9.99	3.03	0.38	9.57	14.18	6.25	6.96	4.15	12.38	13.36	11.65	2.47	1.84	9.95
	max $\epsilon\%$	37.12	37.40	38.55	13.73	1.86	31.44	38.55	20.13	27.74	21.67	37.63	38.55	38.55	8.62	6.66	38.55
Total	ave. π	0.43	0.43	0.36	0.15	0.03	0.23	0.29	0.32	0.27	0.26	0.29	0.32	0.29	0.26	0.27	0.28
	std. π	0.10	0.10	0.12	0.14	0.04	0.20	0.20	0.18	0.19	0.16	0.22	0.25	0.22	0.16	0.15	0.19
	max π	0.64	0.64	0.64	0.51	0.13	0.64	0.64	0.64	0.64	0.43	0.57	0.64	0.64	0.43	0.43	0.64
Total	ave. $\epsilon\%$	4.25	4.29	3.32	0.82	0.08	2.47	3.21	2.14	2.39	1.42	4.13	4.38	3.61	1.12	0.82	2.55
	std. $\epsilon\%$	7.33	7.38	6.03	1.71	0.20	5.63	7.80	4.01	4.59	2.70	7.78	8.43	7.07	1.74	1.29	5.71
	max $\epsilon\%$	37.12	37.40	38.55	13.73	1.86	31.44	38.55	20.13	27.74	21.67	37.63	38.55	38.55	8.62	6.66	38.55
Total	ave. π	0.26	0.26	0.22	0.08	0.01	0.16	0.17	0.17	0.16	0.16	0.17	0.17	0.16	0.16	0.16	0.16
	std. π	0.21	0.21	0.19	0.10	0.02	0.19	0.20	0.19	0.19	0.17	0.21	0.23	0.21	0.17	0.17	0.19
	max π	0.64	0.64	0.64	0.51	0.13	0.64	0.64	0.64	0.64	0.49	0.62	0.64	0.64	0.49	0.49	0.64

Table 4 The summary of the results from scenarios with identical retailers for different h_i levels

h_i		p_i				l_0			l_i			Total
		4	9	19	99	1	3	5	1	3	5	
0.00	<i>ave. $\epsilon\%$</i>	4.56	5.02	3.15	4.27	2.03	7.14	8.02	6.21	1.54	1.07	4.25
	<i>std. $\epsilon\%$</i>	7.82	9.52	4.75	6.35	3.15	9.52	10.62	8.84	1.93	1.39	7.33
	<i>max $\epsilon\%$</i>	31.13	37.12	19.61	27.63	20.76	36.22	37.12	37.12	7.60	5.73	37.12
	<i>ave. π</i>	0.26	0.26	0.26	0.26	0.23	0.29	0.31	0.28	0.23	0.23	0.26
	<i>std. π</i>	0.21	0.21	0.21	0.21	0.18	0.23	0.25	0.23	0.18	0.18	0.21
	<i>max π</i>	0.64	0.64	0.64	0.64	0.49	0.62	0.64	0.64	0.49	0.49	0.64
0.10	<i>ave. $\epsilon\%$</i>	4.63	5.07	3.17	4.29	2.06	7.20	8.05	6.26	1.57	1.10	4.29
	<i>std. $\epsilon\%$</i>	7.90	9.58	4.77	6.35	3.18	9.60	10.69	8.90	1.96	1.43	7.38
	<i>max $\epsilon\%$</i>	31.44	37.40	19.71	27.74	20.94	36.49	37.40	37.40	7.78	5.89	37.40
	<i>ave. π</i>	0.26	0.26	0.26	0.25	0.23	0.29	0.30	0.27	0.23	0.23	0.26
	<i>std. π</i>	0.21	0.21	0.21	0.21	0.18	0.23	0.25	0.23	0.18	0.18	0.21
	<i>max π</i>	0.64	0.64	0.64	0.64	0.49	0.62	0.64	0.64	0.49	0.49	0.64
0.50	<i>ave. $\epsilon\%$</i>	2.66	5.09	2.88	2.66	2.18	5.17	4.90	4.55	1.71	1.23	3.32
	<i>std. $\epsilon\%$</i>	4.08	9.83	4.49	3.00	3.30	8.41	8.38	7.39	2.10	1.57	6.03
	<i>max $\epsilon\%$</i>	15.28	38.55	20.13	9.30	21.67	37.63	38.55	38.55	8.62	6.66	38.55
	<i>ave. π</i>	0.21	0.24	0.23	0.20	0.23	0.21	0.19	0.21	0.23	0.23	0.22
	<i>std. π</i>	0.19	0.21	0.20	0.17	0.18	0.20	0.20	0.20	0.18	0.18	0.19
	<i>max π</i>	0.63	0.64	0.64	0.49	0.49	0.62	0.64	0.64	0.49	0.49	0.64
0.90	<i>ave. $\epsilon\%$</i>	0.45	0.80	1.38	0.62	0.75	1.00	0.82	0.91	0.70	0.66	0.82
	<i>std. $\epsilon\%$</i>	0.69	1.31	2.89	0.89	1.36	2.50	1.70	2.03	1.11	1.01	1.71
	<i>max $\epsilon\%$</i>	3.64	5.39	13.73	4.64	9.00	13.73	8.69	13.73	5.00	4.64	13.73
	<i>ave. π</i>	0.06	0.09	0.09	0.06	0.09	0.05	0.05	0.06	0.10	0.11	0.08
	<i>std. π</i>	0.09	0.12	0.12	0.08	0.11	0.09	0.08	0.09	0.12	0.12	0.10
	<i>max π</i>	0.47	0.43	0.51	0.43	0.47	0.51	0.42	0.51	0.47	0.43	0.51
0.99	<i>ave. $\epsilon\%$</i>	0.03	0.07	0.13	0.10	0.06	0.13	0.11	0.10	0.07	0.06	0.08
	<i>std. $\epsilon\%$</i>	0.03	0.20	0.31	0.10	0.14	0.28	0.23	0.21	0.22	0.10	0.20
	<i>max $\epsilon\%$</i>	0.13	1.86	1.48	0.46	1.86	1.48	1.30	1.48	1.86	0.46	1.86
	<i>ave. π</i>	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.01
	<i>std. π</i>	0.02	0.02	0.03	0.02	0.02	0.02	0.01	0.02	0.03	0.02	0.02
	<i>max π</i>	0.09	0.13	0.13	0.12	0.13	0.09	0.09	0.13	0.13	0.12	0.13
Total	<i>ave. $\epsilon\%$</i>	2.47	3.21	2.14	2.39	1.42	4.13	4.38	3.61	1.12	0.82	2.55
	<i>std. $\epsilon\%$</i>	5.63	7.80	4.01	4.59	2.70	7.78	8.43	7.07	1.74	1.29	5.71
	<i>max $\epsilon\%$</i>	31.44	38.55	20.13	27.74	21.67	37.63	38.55	38.55	8.62	6.66	38.55
	<i>ave. π</i>	0.16	0.17	0.17	0.16	0.16	0.17	0.17	0.16	0.16	0.16	0.16
	<i>std. π</i>	0.19	0.20	0.19	0.19	0.17	0.21	0.23	0.21	0.17	0.17	0.19
	<i>max π</i>	0.64	0.64	0.64	0.64	0.49	0.62	0.64	0.64	0.49	0.49	0.64

probability of imbalance does not necessarily yield a high $\epsilon\%$, but a high π is a necessity for a large relative gap. Among all the instances, 273 and 85 of them have $\epsilon\% < 2$ and $\pi > 0.15$, and $\epsilon\% < 1$ and $\pi > 0.20$, respectively. Second,

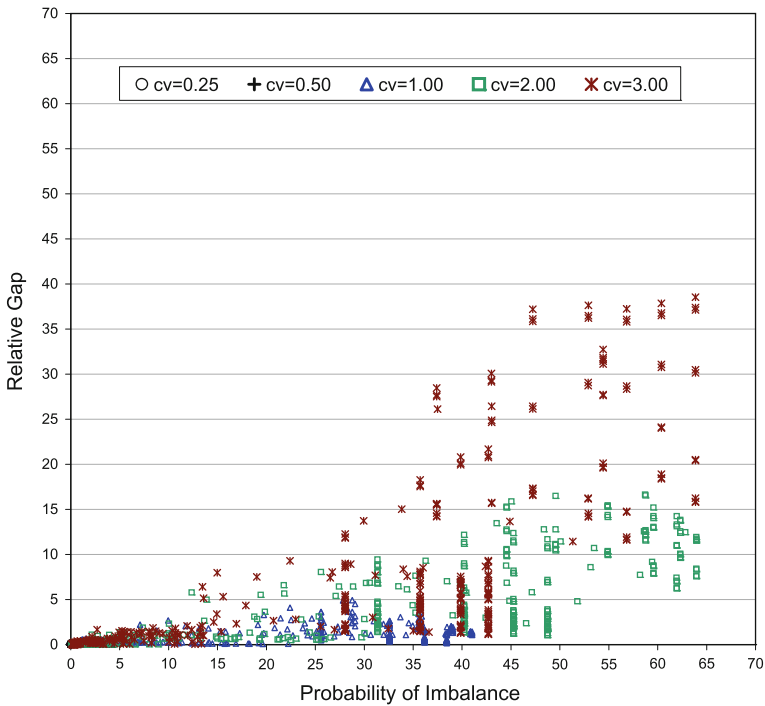


Fig. 2 π (in percentage) vs. $\epsilon\%$ for scenarios with identical retailers

when the points with very high $\epsilon\%$ and π values are analyzed, it turns out that they have long warehouse lead time, short retailer lead time and high coefficient of variation in common. There are 44 points in the region $\epsilon\% \geq 24.03$ and $\pi \geq 0.37$, all coming from scenarios with $cv_i = 2$ or 3 , and $(l_0, l_i) = (3, 1)$ or $(5, 1)$.

4. The probability of having imbalanced retailer inventories tend to increase as the added value at the warehouse increases For any level of another factor (e.g., $l_0 = 1$), observe that *ave. π* and *max π* increases as h_i decreases in Tables 3 and 4.

Having highlighted the main points, next, we further discuss the effect of the coefficient of variation, lead times, and cost parameters on the performance measures.

4.1.1 Coefficient of variation

There is a clear relationship between the coefficient of variation and the imbalance of inventories. For any level of another factor, observe in Table 3 that *max $\epsilon\%$* and *ave. $\epsilon\%$* both grow as the coefficient of variation increases. This phenomenon is intuitive suggesting that as the variance in the demand process gets larger, the

retailer inventories become more imbalanced. $max \pi$ and $ave. \pi$ figures seem to support this argument. In Table 3, for $max \pi$ and $ave. \pi$, a steady rise is observed as the coefficient of variation increases from 0.25 to 2, but the figures drop for 3. In order to explain this decline, we compared the order-up-to level of each of the 400 scenarios with $cv_i = 2$ against the corresponding one with $cv_i = 3$. It turns out that more stock is kept at the warehouse in the scenarios with $cv_i = 3$, which results in lower imbalance probability. Even though the imbalance probability figures decrease for $cv_i = 3$, it does not seem to effect the trend for gap measures.

4.1.2 Holding costs

In order to see the effect of the holding costs better, we extended the study for more h_i values between 0.1 and 0.8, see Table 5. The cost figures in Table 5 are chosen in a way that at one extreme there is no added value at the retailers ($h_i = 0$); at the other extreme, 99% of an item’s value is attained at the retailers ($h_i = 0.99$). Although Table 5 is organized such that the results are summarized with respect to h_i level, the trends observed are also valid when the level of cv_i, p_i, l_0 or l_i is fixed like in Table 4. Our observations are as follows:

- Compare $max \pi$ figures for different h_i values in Table 5. Between 0 and 0.7, $max \pi$ is level at 0.64, and $max \pi$ decreases as h_i grows starting from $h_i = 0.8$. For $ave. \pi$, there is a steady decline as h_i increases. The trend in imbalance probability measures is due to keeping more stock at the warehouse than at the retailers as h_0 decreases (compared to h_i) since it is less costly. As a result, the chance of imbalance decreases.
- Compare $max \epsilon\%$ figures in Table 5. There is an increasing trend starting from $h_i = 0$ until $h_i = 0.5$. Then a sharp decline occurs at $h_i = 0.6$ ($max \epsilon\% = 38.55$ for $h_i = 0.5$ and $max \epsilon\% = 20.15$ for $h_i = 0.6$). From this point on, $max \epsilon\%$ values decrease with another sharp decline at $h_i = 0.99$. For $ave. \epsilon\%$, there is a steady decreasing trend as h_i increases starting from $h_i = 0.1$ with sharp declines at $h_i = 0.6$ and $h_i = 0.99$.

These observations point out a general tendency that the relative gap decreases as the added value at the upper echelon decreases relative to the one at the lower

Table 5 The summary of the results for various values of h_i ($= 1 - h_0$) with identical retailers

	h_i										
	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.99
<i>ave. $\epsilon\%$</i>	4.25	4.29	4.24	3.75	3.50	3.32	2.38	1.88	1.41	0.82	0.08
<i>std. $\epsilon\%$</i>	7.33	7.38	7.28	6.32	6.12	6.03	3.45	2.81	2.34	1.71	0.20
<i>max $\epsilon\%$</i>	37.12	37.40	37.68	37.96	38.25	38.55	20.15	18.99	17.30	13.73	1.86
<i>ave. π</i>	0.26	0.26	0.25	0.24	0.23	0.22	0.19	0.16	0.12	0.08	0.01
<i>std. π</i>	0.21	0.21	0.21	0.20	0.19	0.19	0.18	0.16	0.14	0.10	0.02
<i>max π</i>	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.64	0.59	0.51	0.13

echelon. This pattern is more apparent when the added value at the warehouse is less than the one at the retailers (i.e., for $h_i > 0.5$).

4.1.3 Lead times

Recall from the analysis of variance that the main effect of l_0 as well as its interactions with the coefficient of variation or the holding cost are found to be significant. The overall trend observed in Tables 3 and 4 is that all statistics tend to increase as l_0 grows. For example, average (standard deviation) of relative gap is 1.42 (2.70), 4.13 (7.78) and 4.38 (8.43) for l_0 equal to 1, 3 and 5, respectively. The logic behind this behavior is that it takes a longer time to recover from an imbalance situation when the warehouse lead time is long. If the system stays in imbalance for more periods in a row, the effect on the expected cost amplifies. However, this behavior seem not to hold for high h_i or low cv_i . As an example, consider $h_i = 0.9$ in Table 4. All statistics increase when the warehouse lead time is extended from 1 to 3, but the same statistics decrease when l_0 is further increased to 5.

Similar to the warehouse lead time, the main effect of the retailer lead time as well as its interactions with the coefficient of variation or the holding cost are found to be significant in the analysis of variance. Table 3 (4) show that for any level of cv_i (h_i) all relative gap statistics decrease as l_i increases.

In the light of the trends observed for lead times, we extend the numerical study for more retailer lead time values while fixing the warehouse lead time at 1. The results are tabulated in Table 6. The first observation is that the relative gap is *small* when the warehouse lead time is short, and retailer lead time is long (i.e., $(l_0, l_i) = (1, 7), (1,9), (1,11)$ or $(1,13)$). This observation is valid even if cv_i or h_i is fixed at any level considered in the numerical study. The steady decline in *ave. $\epsilon\%$* and *std. $\epsilon\%$* and a decreasing trend in *max $\epsilon\%$* in Table 6 indicate that the relative gap shrinks as retailer lead time grows. This observation might seem counterintuitive at first sight because the variance faced by the system grows and the warehouse is forced to ration far in advance as retailer lead time increases. This is expected to have a negative effect on the balance of retailers and reflected by flat imbalance probability statistics for $l_i \geq 3$. However, the impact of retailer lead time on the mean and the standard deviation of demand is different resulting in a lower

Table 6 The summary of the results for l_i with identical retailers

	(l_0, l_i)						
	(1,1)	(1,3)	(1,5)	(1,7)	(1,9)	(1,11)	(1,13)
<i>ave. $\epsilon\%$</i>	3.61	1.12	0.82	0.60	0.47	0.43	0.38
<i>std. $\epsilon\%$</i>	7.07	1.74	1.29	0.77	0.63	0.63	0.55
<i>max $\epsilon\%$</i>	38.55	8.62	6.66	3.08	2.83	3.01	2.19
<i>ave. π</i>	0.16	0.16	0.16	0.17	0.16	0.17	0.17
<i>std. π</i>	0.21	0.17	0.17	0.17	0.17	0.17	0.17
<i>max π</i>	0.64	0.49	0.49	0.49	0.49	0.49	0.49

coefficient of variation. As l_i increases, the mean and standard deviation of demand over the retailer lead time increases at the respective rate of l_i and $\sqrt{l_i}$, which yields a coefficient of variation decreasing at a rate $\sqrt{l_i}$. As we have demonstrated clearly, low coefficient of variation leads to smaller relative gaps.

4.1.4 Penalty costs

Even though the effect of the penalty cost is found to be statistically significant in the analysis of variance, the results do not show a clear trend between p_i and the measures considered. Consider the overall $\epsilon\%$ statistics in Table 3 in row titled "Total". The measures first increase, then decrease and finally increase again as p_i increases. Although $ave.\epsilon\%$, $std.\epsilon\%$ and $max\epsilon\%$ show an increasing trend for $cv_i = 0.25, 0.5, 1$, the results are mixed for $cv_i = 2, 3$. Moreover, we do not observe a common trend for each level of h_i in Table 4. Thus, we conclude that the results do not reveal any clear trend for the relative gap as p_i is perturbed.

4.2 Nonidentical retailers

We restricted the number of retailers to two ($N = 2$) so that the effect of asymmetry in each parameter can be captured easily. The following parameter settings are used:

$$\begin{aligned}(h_0, h_1, h_2) &= (0.5, 0, 0.5), (0.5, 0.5, 0), (0.5, 0.5, 0.5), (0.9, 0, 0.1), \\ &\quad (0.9, 0.1, 0), (0.9, 0.1, 0.1) \\ (cv_1, cv_2) &= (0.5, 0.5), (0.5, 1), (0.5, 2), (1, 0.5), (1, 1), (1, 2), (2, 0.5), (2, 1), (2, 2) \\ (l_0, l_1, l_2) &= (1, 1, 1), (1, 1, 5), (1, 5, 1), (1, 5, 5), (5, 1, 1), (5, 5, 5) \\ (\mu_1, \mu_2) &= (0.05, 0.95), (0.2, 0.8), (0.5, 0.5) \\ (p_1, p_2) &= (9, 9), (9, 19), (19, 9), (19, 19)\end{aligned}$$

All combinations of the parameters generate a test bed of 3888 problem instances. All through the paper, we interpret the mean demand as a representative of the size of a retailer. The values for mean demands are chosen to see the effect of big/small retailer and without loss of generality, the sum is kept at 1. Note that in all the problem instances considered, the first retailer is equal or smaller in size compared to the second. Three values are selected for the coefficient of variation: 0.5 (low), 1 (moderate), 2 (high). This leads to nine distinct combinations. We picked lead time values to reflect the effect of short warehouse and retailer lead times [(1,1,1)], long warehouse and retailer lead times [(5,5,5)], long warehouse-short retailer lead times [(5,1,1)] and vice versa [(1,5,5)], and short warehouse lead time and asymmetric retailer lead times [(1,1,5) and (1,5,1)]. From two values selected for penalty costs, high (19) and moderate (9), four combinations are generated to see the effect of equal and asymmetric penalty costs. We interpret the holding costs as a way of reflecting the added value. While (0.5,0.5,0.5) combination corresponds to an equal added value at all stock points, (0.9,0.1,0.1) represents high added value at the warehouse. The other combinations are chosen to see the effect of having zero added value at one of the retailers. This situation is relevant to production and

distribution environments. Consider a finished product or a semi-finished assembly that is stored centrally to supply two downstream stock points: at one, negligible value is added, e.g., just transportation or packaging, and at the other there is a positive added value, e.g., a module is added or further assembly operations are carried out. In a distribution environment, the added value at the retailers may consist of transportation costs and the cost of the storage space, which might be almost zero for one and considerable for another.

Recall that except the number of retailers, all other parameters in the stochastic model are found to have a significant impact on the relative gap. Even though we restrict the number of retailers to two, the number of factors and their interactions to be considered in an analysis of variance is high in the nonidentical retailers case. Hence, we decided not to conduct an analysis of variance, but to discuss the trends and patterns in the results. A summary of the results can be found in Table 7, which is similar to Tables 3 and 4 in terms of organization. The main findings are given below.

1. Recall that we consider the conditions $\max \epsilon\% \leq 5$ and $\text{ave. } \epsilon\% \leq 1$ as criteria for a relative gap to be small. The results in Table 7 show that for nonidentical retailers, the relative gap is small when the warehouse lead time is short, and the lead time of the second retailer (equal or larger in size) is long (i.e., $(l_0, l_1, l_2) = (1, 5, 5)$ or $(1, 1, 5)$).

There are 1,888, 2,652 and 3,390 problem instances having $\epsilon\%$ less than or equal to 1, 2, and 5, respectively. The instances with high $\epsilon\%$ figures have the following common parameter settings:

- 27 scenarios with $\epsilon\% \geq 69.76$; all having $(l_0, l_1, l_2) = (5, 1, 1)$, $(h_0, h_1, h_2) = (0.5, 0, 0.5)$, and $(\mu_1, \mu_2) = (0.05, 0.95)$,
- 36 scenarios with $69.76 > \epsilon\% \geq 43.65$; all having $h_1 = 0$ and $(l_0, l_1, l_2) = (5, 1, 1)$,
- 55 scenarios with $43.65 > \epsilon\% > 25$; all having $h_1 = 0$ and $l_0 = 5$,
- 23 scenarios with $25 \geq \epsilon\% > 17.85$; all having $h_1 = 0$ or $h_2 = 0$, and $l_0 = 5$,
- 35 scenarios with $17.85 \geq \epsilon\% > 15.3$; all having $h_1 = 0$ or $h_2 = 0$.

The figures above reveal the fact that long warehouse lead time and having zero added value at one of the retailers may yield high gaps (especially zero added value at the small retailer). A thorough explanation of this issue is presented later where holding costs are discussed.

It is interesting to notice that although the coefficient of variation turned out to be the main determinant of gaps in the identical retailers case, even $(cv_1, cv_2) = (0.5, 0.5)$ may lead to $\epsilon\%$ as high as 178.17 in the nonidentical retailers case. Unlike in the identical retailers case, the relative gap is small only in a few parameter settings. In many practically relevant parameter combinations, we observe large gaps. Especially, when there is asymmetry in the size of the retailers, long warehouse lead time, or negligible added value at one of the retailers, large gaps may occur.

2. The Pearson correlation coefficient between $\epsilon\%$ and π is 0.542, which indicates a positive correlation. However, while a high relative gap requires a high imbalance probability, a high imbalance probability does not necessarily correspond to a high gap. The imbalance probabilities are plotted against the corresponding $\epsilon\%$ values in Fig. 3. Observe that no point is in the region low π and high $\epsilon\%$. In the region with high $\epsilon\%$ and π , there are 118 problem instances with $\epsilon\% \geq 25.94$ and $\pi\% \geq 62.51$. All these instances have $l_0 = 5$ and zero added value at the small retailer ($\mu_1 = 0.05$ or 0.2).

In the rest of this subsection, we discuss the relationships between the input parameters and the measures on the basis of trends. For example, we analyze how relative gap and imbalance probability measures are affected when μ_1 is decreased with respect to μ_2 , and identify whether a clear trend exists.

4.2.1 Retailer size

The results show that as the size of the retailers become disproportionate, the relative gap grows. When μ_1 is decreased with respect to μ_2 , we observe that all gap and probability measures increase. The impact of retailer size will be more evident as we discuss the other parameters like holding costs and coefficient of variation.

4.2.2 Holding costs

- Similar to the identical retailers case, we observe that as the added value at the warehouse increases, the retailer inventories tend to become more imbalanced. Compare $\max \epsilon\%$ and $\text{ave. } \epsilon\%$ values in columns (0.5,0.5,0.5) and (0.9,0.1,0.1) of Table 7. When h_0 gets larger, the system prefers to keep less stock at the warehouse, which in return increases the chance of imbalance. Larger $\max \pi$ and $\text{ave. } \pi$ in column (0.9,0.1,0.1) support the claim.
- We can also discuss the combined effect of size and holding costs. Recall that the first retailer is equal to or smaller than the second one with respect to size. In Table 7, compare the following columns: (0.5,0,0.5) to (0.5,0.5,0.5), and (0.9,0,0.1) to (0.9,0.1,0.1). The results indicate that when the added value at the first retailer becomes zero, imbalance emerges as an important issue. Note that low added value at a retailer increases its order-up-to level, see Eq. (7), so instead of retaining stock, the warehouse ships to this retailer. Since little or no inventory is kept at the warehouse, the chance of imbalance increases. There are two observations that verify this argument. First, all measures show a considerable increase when the holding cost at the first retailer becomes zero. Second, in terms of all measures the trend is stronger when $h_0 = 0.5$ because all the stock kept at the warehouse is routed to the small retailer. Since not much stock is kept at the warehouse when $h_0 = 0.9$ (due to high holding cost), the impact is weaker compared to $h_0 = 0.5$ case. We do not observe such a strong trend when the added value at the second retailer decreases; compare the following columns in Table 7: (0.5,0.5,0) to (0.5,0.5,0.5), and (0.9,0.1,0) to (0.9,0.1,0.1).

Table 7 The summary of the results from scenarios with nonidentical retailers, $N = 2$

cv_1		cv_2			Total
		0.5	1	2	
0.5	<i>ave.</i> $\epsilon\%$	3.15	3.24	1.95	2.78
	<i>std.</i> $\epsilon\%$	16.08	14.78	6.94	13.24
	<i>max</i> $\epsilon\%$	178.17	155.96	69.76	178.17
	<i>ave.</i> π	0.12	0.16	0.18	0.15
	<i>std.</i> π	0.19	0.18	0.16	0.18
	<i>max</i> π	0.86	0.90	0.86	0.90
1	<i>ave.</i> $\epsilon\%$	3.92	4.18	2.91	3.67
	<i>std.</i> $\epsilon\%$	16.16	14.79	7.05	13.29
	<i>max</i> $\epsilon\%$	179.57	156.47	69.81	179.57
	<i>ave.</i> π	0.25	0.29	0.28	0.27
	<i>std.</i> π	0.17	0.15	0.14	0.16
	<i>max</i> π	0.86	0.90	0.86	0.90
2	<i>ave.</i> $\epsilon\%$	5.48	5.84	4.56	5.29
	<i>std.</i> $\epsilon\%$	16.82	15.16	7.42	13.76
	<i>max</i> $\epsilon\%$	186.86	159.42	70.23	186.86
	<i>ave.</i> π	0.40	0.43	0.36	0.39
	<i>std.</i> π	0.15	0.13	0.13	0.14
	<i>max</i> π	0.86	0.90	0.86	0.90
Total	<i>ave.</i> $\epsilon\%$	4.25	4.42	3.14	3.91
	<i>std.</i> $\epsilon\%$	7.33	14.94	7.21	13.47
	<i>max</i> $\epsilon\%$	37.12	159.42	70.23	186.86
	<i>ave.</i> π	0.26	0.29	0.27	0.27
	<i>std.</i> π	0.21	0.19	0.16	0.19
	<i>max</i> π	0.64	0.90	0.86	0.90

p_1		p_2		Total
		9	19	
9	<i>ave.</i> $\epsilon\%$	3.33	4.75	4.04
	<i>std.</i> $\epsilon\%$	11.19	15.56	13.57
	<i>max</i> $\epsilon\%$	129.53	186.86	186.86
	<i>ave.</i> π	0.27	0.28	0.28
	<i>std.</i> π	0.18	0.19	0.19
	<i>max</i> π	0.90	0.90	0.90
19	<i>ave.</i> $\epsilon\%$	3.34	4.23	3.79
	<i>std.</i> $\epsilon\%$	11.13	15.27	13.36
	<i>max</i> $\epsilon\%$	127.00	182.49	182.49
	<i>ave.</i> π	0.27	0.27	0.27
	<i>std.</i> π	0.18	0.19	0.19
	<i>max</i> π	0.90	0.90	0.90

Table 7 continued

p_1	p_2		Total			
			9	19		
Total	<i>ave. ε%</i>		3.33	4.49	3.91	
	<i>std. ε%</i>		11.16	15.41	13.47	
	<i>max ε%</i>		129.53	186.86	186.86	
	<i>ave. π</i>		0.27	0.27	0.27	
	<i>std. π</i>		0.18	0.19	0.19	
	<i>max π</i>		0.90	0.90	0.90	
<hr/>						
(h_0, h_1, h_2)						
<hr/>						
	(0.5,0,0.5)	(0.5,0.5,0)	(0.5,0.5,0.5)	(0.9,0,0.1)	(0.9,0.1,0)	(0.9,0.1,0.1)
<i>ave. ε%</i>	13.71	1.73	1.28	3.72	1.48	1.56
<i>std. ε%</i>	29.56	2.86	1.49	8.79	1.98	2.07
<i>max ε%</i>	186.86	24.76	8.12	67.90	15.35	14.53
<i>ave. π</i>	0.43	0.20	0.19	0.32	0.24	0.25
<i>std. π</i>	0.24	0.12	0.12	0.20	0.14	0.14
<i>max π</i>	0.90	0.50	0.51	0.89	0.58	0.60
<hr/>						
(l_0, l_1, l_2)						
<hr/>						
	(1,1,1)	(1,1,5)	(1,5,1)	(1,5,5)	(5,1,1)	(5,5,5)
<i>ave. ε%</i>	2.43	0.81	2.19	0.59	13.98	3.48
<i>std. ε%</i>	2.58	0.96	2.34	0.64	29.86	7.53
<i>max ε%</i>	17.84	4.90	17.32	2.93	186.86	43.30
<i>ave. π</i>	0.25	0.22	0.27	0.25	0.32	0.32
<i>std. π</i>	0.16	0.13	0.16	0.14	0.25	0.22
<i>max π</i>	0.75	0.67	0.74	0.68	0.90	0.90
<hr/>						
(μ_1, μ_2)						
<hr/>						
	(0.05,0.95)		(0.2,0.8)		(0.5,0.5)	
<i>ave. ε%</i>	6.12		3.25		2.37	
<i>std. ε%</i>	21.72		7.46		3.02	
<i>max ε%</i>	186.86		69.66		24.76	
<i>ave. π</i>	0.31		0.27		0.24	
<i>std. π</i>	0.25		0.16		0.12	
<i>max π</i>	0.90		0.71		0.50	

Table 8 shows all scenarios with (h_0, h_1, h_2) equal to $(0.5,0,0.5)$ and $(0.5,0.5,0)$, grouped with respect to the mean demands. When the added value at the first retailer is zero, all statistics show a sharp increase when the first retailer’s mean demand decreases. In contrast, the aforementioned measures except *max π* and *std. π* decrease as the second retailer’s mean demand increases in case $(0.5,0.5,0)$.

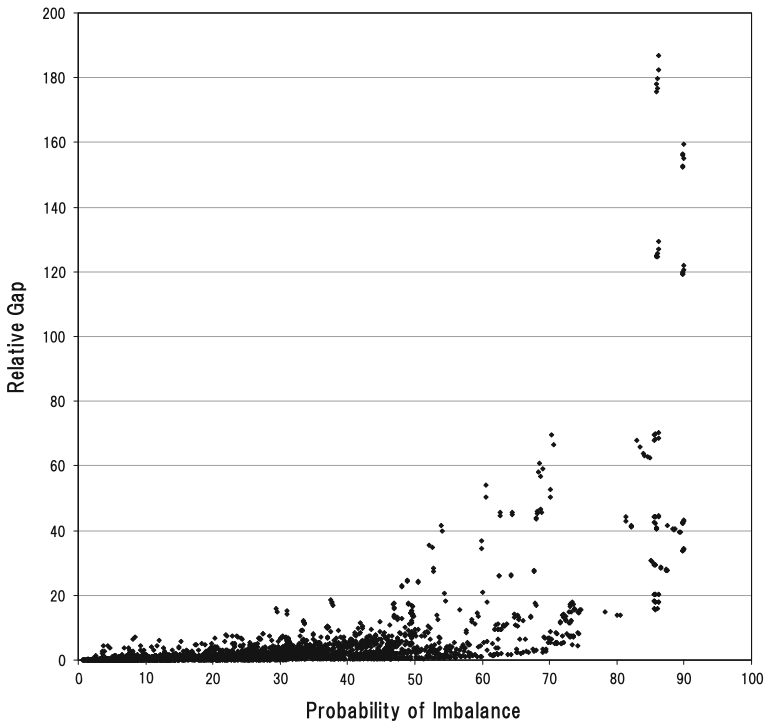


Fig. 3 π (in percentage) vs. $\epsilon\%$ for scenarios with nonidentical retailers

Table 8 Forwarding-to-the-small-retailer effect

(μ_1, μ_2)	(h_0, h_1, h_2)					
	$(0.5, 0, 0.5)$			$(0.5, 0.5, 0)$		
	$(0.05, 0.95)$	$(0.2, 0.8)$	$(0.5, 0.5)$	$(0.5, 0.5)$	$(0.2, 0.8)$	$(0.05, 0.95)$
<i>ave. $\epsilon\%$</i>	27.17	10.63	3.33	3.39	1.25	0.55
<i>std. $\epsilon\%$</i>	45.54	15.49	4.09	4.15	1.47	0.89
<i>max $\epsilon\%$</i>	186.86	69.66	24.26	24.76	8.89	5.37
<i>ave. π</i>	0.62	0.42	0.26	0.26	0.19	0.15
<i>std. π</i>	0.26	0.18	0.11	0.11	0.11	0.12
<i>max π</i>	0.90	0.71	0.50	0.50	0.40	0.43

(Note that the table is organized in such a way that the mean demand of the retailer with zero additional inventory cost increases from left to right.)

In the light of these observations, we conclude that when the added value at one retailer is zero, both the probability and the extent of imbalance grows considerably as this retailer’s mean demand decreases with respect to the other. Under the balance assumption, the amount of stock to keep at the warehouse is underestimated. Since the optimal order-up-to level for the retailer with zero

added value is infinity, anything that might have been kept at the warehouse is sent to this retailer. As the mean demand of this retailer decreases it creates more and larger imbalance in the system. We call this phenomenon of the system getting severely imbalanced due to no added value at the small retailer as *forwarding-to-the-small-retailer*.

4.2.3 Penalty costs

Unlike the identical retailers case, interesting relationships are identified:

- The results confirm that as the penalty costs get larger, the relative gap grows. When (9,9) is compared against (19,19) in Table 7, all relative gap statistics increase.
- For the effect of size, compare: i. (9,9) against (9,19), ii. (9,9) against (19,9). In case (i), while probability measures do not show a trend, $max\ \epsilon\%$, $ave.\ \epsilon\%$ and $std.\ \epsilon\%$ all increase when the penalty cost of the second retailer is increased. However, in case (ii), there is hardly any change in any of the measures. We can conclude that as the penalty cost of the second retailer increases, the relative gap widens. When the penalty cost of the big retailer is high, the negative effect of overstocking at the small retailer gets stronger, resulting in a larger gap.

4.2.4 Coefficient of variation

In line with the identical retailers case, the impact of the coefficient of variation on the imbalance of inventories is well established. Compare the statistics for (0.5,0.5), (1,1) and (2,2) in Table 7. We observe that $ave.\ \epsilon\%$ grows while $std.\ \epsilon\%$ shrinks. Interestingly, $max\ \epsilon\%$ shows a considerable decrease as the demand variance grows. All these indicate an increasing trend in the relative gap as the coefficient of variation increases. Growing $ave.\ \pi$ figures also support this tendency.

The joint effect of the retailer size and the coefficient of variation is more complicated. Recall that the first retailer is smaller or equal to the second retailer in terms of size (mean demand). In Table 7, when the second retailer's coefficient of variation is kept constant and the first retailer's is varied (compare columns i. (0.5,0.5), (1,0.5), (2,0.5) ii. (0.5,1), (1,1), (2,1) iii. (0.5,2), (1,2), (2,2)), or the coefficient of variation of the first retailer is fixed and the second retailer's is varied (compare columns iv. (0.5,0.5), (0.5,1), (0.5,2) v. (1,0.5), (1,1), (1,2) vi. (2,0.5), (2,1), (2,2)), we observe different trends in the measures. In cases (i)–(iii), all measures show a positive trend as the coefficient of variation of the first retailer increases. For example, in case (i), $ave.\ \epsilon\%$ is 3.15, 3.92 and 5.48 for (0.5,0.5), (1,0.5) and (2,0.5), respectively. This behavior is intuitive since the demand variability faced by the system grows. However, the results are mixed for cases (iv)–(vi). On one hand, $max.\ \epsilon\%$ figures decrease considerably as cv_2 is increased. On the other hand, $ave.\ \epsilon\%$ first increases (as cv_2 increases to 1), and then decreases when cv_2 becomes 2. For example, in case (iv), $ave.\ \epsilon\%$ is 3.15, 3.24 and 1.95 for (0.5,0.5), (0.5,1) and (0.5,2), respectively. These results hint that size asymmetry

and coefficient of variation have a more complicated joint effect on the balance of inventories. Hence, we conducted further analysis, see Dođru et al. (2005), but the results had not revealed any clear trend or findings.

4.2.5 Lead times

- *Warehouse lead time:* The impact of warehouse lead time is in line with the observations from the identical retailers case. When the columns (1,1,1) and (5,1,1) of Table 7 are compared, one sees the drastic effect of warehouse lead time: $max \epsilon\%$, $ave. \epsilon\%$ and $std. \epsilon\%$ obtain very high values for $l_0 = 5$. When the warehouse lead time extends, the probability of having imbalanced inventories increases and an imbalance situation takes a longer time to fix, both of which widen the gaps.
- *Retailer lead times:* In Table 7, compare column (1,1,1) to (1,5,5). Although imbalance probability measures do not show a clear difference, $max \epsilon\%$, $ave. \epsilon\%$ and $std. \epsilon\%$ figures point out a considerable decrease as both retailer lead times increase. This observation is in line with the results of the identical retailers case. Next, we consider the impact of asymmetry in retailer lead times. The relative gap statistics are lower for (1,1,5) and (1,5,1) in comparison to (1,1,1). The extent of the decrease seems to depend on the size of the retailer whose lead time grows. If the big retailer's lead time is increased (compare (1,1,5) against (1,1,1)), then the decrease in the relative gap tend to be more in comparison to the decrease observed when the small retailer's lead time is increased (compare (1,5,1) against (1,1,1)). This observation can be explained by the phenomenon discussed for the identical retailers case that the lower coefficient of variation over the retailer lead time results in the gap shrinking. The decrease in the coefficient of variation of demand over the lead time of the retailer with the lower mean demand has limited impact.

4.3 Summary and insights

In this subsection, we summarize the results and insights obtained. Recall that a setting is considered to lead to a small relative gap when $max \epsilon\% \leq 5$ and $ave. \epsilon\% \leq 1$.

The results from the test bed of 2000 scenarios for the identical retailers show that the relative gap is small (irrespective of the values of the other input parameters) when one of the following conditions is satisfied:

- the coefficient of variation is low or moderate ($cv_i = 0.25$ or 0.5),
- the added value at the warehouse is insignificant with respect to the retailers ($h_0 = 0.01$ and $h_i = 0.99$),
- short warehouse lead time and long retailer lead times ($(l_0, l_i) = (1, 7), (1, 9), (1, 11)$ or $(1, 13)$).

Under these settings, the use of the balance assumption is justified. Especially when the coefficient of variation is low (i.e., $cv_i = 0.25$ or 0.5), $max. \epsilon\% \leq 0.36$ and

ave. $\epsilon\%$ is negligible. If none of these conditions hold then high relative gaps may occur; we have observed $\epsilon\% = 38.6$. A high coefficient of variation, a long warehouse lead time, and a low or moderate added value at the retailers (relative to the warehouse) are the main determinants of large relative gaps.

The results from the test bed of 3,888 problem instances with nonidentical retailers indicate that the relative gap is small when the warehouse lead time is short and the lead time of the big retailer is long ($(l_0, l_1, l_2) = (1, 1, 5)$ or $(1, 5, 5)$). Note that the relative gap is small over a limited set of input parameters when the retailers are not identical. Large gaps (as high as 186.9%) are observed in scenarios with long warehouse lead time and/or positive added value at one retailer and zero at the other. The last point regarding negligible added values has also been reported by Gallego et al. (2007).

In the test bed for identical retailers, it is observed that the relative gap increases as (1) the added value at the warehouse (h_0) increases relative to the one at the retailers (h_i), (2) the coefficient of variation (cv_i) increases, (3) the warehouse lead time (l_0) increases, (4) the retailer lead times (l_i) decreases. The last point is an interesting observation because as the retailer lead times extend, the demand variance faced by the system increases and the warehouse rations the on-hand stock far in advance, which may seem to cause the relative gap to grow. However, as the retailer lead times get larger, the coefficient of variation of demand over the lead time decreases, which leads to smaller relative gaps. The aforementioned four points are also observed in the test bed for nonidentical retailers. In addition, it is observed that the relative gap increases as (5) the asymmetry in the size of the retailers grows (i.e., as μ_1 and μ_2 become more disproportionate), (6) the added value at the small retailer becomes zero, (7) the size (i.e., the mean demand) of the retailer with zero added value decreases, (8) the penalty cost of the big retailer (p_2) increases. The phenomenon in (6) is referred to as *forwarding-to-the-small-retailer*, which is a result of overstocking at the small retailer due to negligible added value at this retailer. Forwarding-to-the-small-retailer leads to a severely imbalanced system, which results in big gaps. In such a case, one may consider putting an upper bound on the order-up-to level of the small retailer.

Scenarios with a high relative gap also exhibits a high imbalance probability, but a high imbalance probability does not necessarily imply a large relative gap. There are 85 and 601 scenarios with $\epsilon\% < 1$ and $\pi > 20$ in identical and nonidentical retailers cases, respectively. These results may be important for developing heuristics that are based on or use imbalance probability (e.g., Verrijdt and de Kok (1995) and van der Heijden (1997)). Eppen and Schrage (1981) derived an approximate term for the probability of having a balanced allocation, and the findings reveal the fact that probability of imbalance grows as the number of retailers increases; our results for the imbalance probabilities are in line with this fact. In addition, our results show that the imbalance probability increases as the warehouse lead time extends, the added value at the warehouse increases (with respect to the one at the retailers), and the retailers become more disproportionate in terms of size.

5 Conclusion and further research

In this paper, we studied the effect of the balance assumption on the average expected cost in divergent inventory systems. The balance assumption is widely used in the analysis of periodic review divergent inventory systems and accepted to lead to a good approximation in many different settings. A relaxation of the original problem is obtained under the balance assumption, and the corresponding average expected cost is a lower bound. When the optimal policy of the relaxed model is modified (LB heuristic policy) and simulated, an upper bound is obtained. The relative gap between these bounds is used as a measure for the effect of the balance assumption. We explicitly determined the scenarios that lead to small relative gaps (i.e., the settings under which the use of the balance assumption is justified). In many practically relevant scenarios, which are identified expressly, the relative gaps are found to be moderate or large. This implies either the lower bound or the LB heuristic policy, or both are mediocre for such settings. For future research, the original optimization problem can be solved optimally by dynamic programming. Due to the curse of dimensionality, it is not possible to solve the problem instances of this paper (even for the simple one-warehouse two-retailers system) as they are. Thus, simple discrete demand structures and limited input parameters can only be considered. The results may help to assess the precise impact of the balance assumption and shed light on the optimal policy behavior, which can be important for developing better heuristics. We are actively working on this follow-up study.

Acknowledgments The authors would like to thank two anonymous referees for their helpful comments. This research was completed while Mustafa K. Dođru was a PhD student at Technische Universiteit Eindhoven, and he was supported by The Netherlands Organization for Scientific Research (NWO) with grant number 425-10-004. Mustafa K. Dođru was partly supported by Industrial Development Agency (IDA), Ireland.

References

- Aviv Y, Federgruen A (2001) Capacitated multi-item inventory systems with random and seasonally fluctuating demands: implications for postponement strategies. *Manage Sci* 47:512–531
- Axsäter S (2003) Supply chain operations: serial and distribution inventory systems. In: de Kok AG, Graves SC (eds) *Handbook in operations research and management science*, volume 11: design and analysis of supply chains. Elsevier, Amsterdam, pp 525–559
- Axsäter S, Marklund J, Silver EA (2002) Heuristic methods for centralized control of one-warehouse, N -retailer inventory systems. *M&SOM* 4:75–97
- Bollapragada S, Akella R, Srinivasan R (1998) Centralized ordering and allocation policies in a two-echelon system with non-identical warehouses. *Eur J Oper Res* 106:74–82
- Cachon GP, Fisher M (2000) Supply chain inventory management and the value of shared information. *Manage Sci* 46:1032–1048
- Cao D, Silver EA (2005) A dynamic allocation heuristic for centralized safety stock. *Naval Res Log* 52:513–526
- Chen F, Zheng YF (1994) Lower bounds for multi-echelon stochastic inventory problems. *Manage Sci* 40:1426–1443
- Clark AJ, Scarf H (1960) Optimal policies for a multi-echelon inventory problem. *Manage Sci* 6:475–490
- de Kok AG (1990) Hierarchical production planning for consumer goods. *Eur J Oper Res* 45:55–69
- Diks EB, de Kok AG (1998) optimal control of a divergent N -echelon inventory system. *Eur J Oper Res* 111:75–97

- Diks EB, de Kok AG (1999) Computational results for the control of a divergent N -echelon inventory system. *Int J Prod Econ* 59:327–336
- Dođru MK (2006) Optimal control of one-warehouse multi-retailer systems: an assessment of the balance assumption. Ph.D. thesis, Beta Research School, D77, Technische Universiteit Eindhoven, Eindhoven
- Dođru MK, van Houtum GJ, de Kok AG (2005) A numerical study on the effect of the balance assumption in one-warehouse multi-retailer inventory systems. Beta Working Paper, WP 135, Department of Technology Management, Technische Universiteit Eindhoven
- Eppen G, Schrage L (1981) Centralized ordering policies in a multi-warehouse system with lead times and random demand. In: Schwarz LB (ed) *Multi-level production-inventory control systems: theory and practice*. North-Holland, pp 51–67
- Erkip N, Hausman WH, Nahmias S (1990) Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands. *Manage Sci* 36:381–392
- Federgruen A, Zipkin P (1984a) Allocation policies and cost approximations for multilocation inventory systems. *Naval Res Log Q* 31:97–129
- Federgruen A, Zipkin P (1984b) Approximations of dynamic, multi-location production and inventory problems. *Manage Sci* 30:69–84
- Federgruen A, Zipkin P (1984c) Computational issues in an infinite-horizon, multi-echelon inventory model. *Oper Res* 32:818–836
- Gallego G, Özer Ö, Zipkin P (2007) Bounds, heuristics, and approximations for distribution systems. *Oper Res* 55:503–517
- Graves SC (1996) A multiechelon inventory model with fixed replenishment intervals. *Manage Sci* 42:1–18
- Jackson PL (1988) Stock allocation in a two-echelon distribution system or “what to do until your ship comes in”. *Manage Sci* 34:880–895
- Jackson PL, Muckstadt JA (1989) Risk pooling in a two-period, two-echelon inventory stocking and allocation problem. *Naval Res Log* 36:1–26
- Jönsson H, Silver EA (1987) Analysis of a two-echelon inventory control system with complete redistribution. *Manage Sci* 33:215–277
- Kumar A, Jacobson SH (1998) Optimal and near-optimal decisions for procurement and allocation of a critical resource with a stochastic consumption rate. Working Paper, University of Michigan Business School
- Kumar A, Schwarz LB, Ward JE (1995) Risk-pooling along a fixed delivery route using a dynamic inventory-allocation policy. *Manage Sci* 41:344–362
- Kunnunkal S, Topaloglu H (2008) A duality-based relaxation and decomposition approach for inventory distribution systems. *Naval Res Log* 55:612–631
- McGavin EJ, Schwarz LB, Ward JE (1993) Two-interval inventory-allocation policies in a one-warehouse, N -identical-retailer distribution system. *Manage Sci* 39:1092–1107
- Özer Ö (2003) Replenishment strategies for distribution systems under advance demand information. *Manage Sci* 49:255–272
- Tijms HC (2003) *A first course in stochastic models*. Wiley, Chichester
- van der Heijden MC (1997) Supply rationing in multi-echelon divergent systems. *Euro J Oper Res* 101:532–549
- van der Heijden MC (1999) Multi-echelon inventory control in divergent systems with shipping frequencies. *Euro J Oper Res* 116:331–351
- van der Heijden MC, Diks EB, de Kok AG (1997) Stock allocation in general multi-echelon distribution systems with (R, S) order-up-to-policies. *Int J Prod Econ* 49:157–174
- van Donselaar K (1990) Integral stock norms in divergent systems with lot-sizes. *Euro J Oper Res* 45:70–84
- van Donselaar K, Wijngaard J (1987) Commonality and safety stocks. *Eng Costs Prod Econ* 12:197–204
- Verrijdt JHCM, de Kok AG (1995) Distribution planning for a divergent N -echelon network without intermediate stocks under service restrictions. *Int J Prod Econ* 38:225–243
- Zipkin P (1984) On the imbalance of inventories in multi-echelon systems. *Math Oper Res* 9:402–423

Author Biographies

Mustafa K. Dođru is a member of the technical staff in the Industrial Mathematics and Operations Research department of Alcatel-Lucent Bell Labs in Dublin, Ireland. He received B.S. and M.S. degrees in industrial engineering from the Middle East Technical University, Ankara, Turkey, and a Ph.D. in OR/OM from Technische Universiteit Eindhoven, the Netherlands. His current research interests include multi-echelon inventory systems, assemble-to-order production/inventory models, and multi-agent negotiations.

A. G. de Kok received his B.Sc. in Mathematics and Physics in 1979 and his M.Sc. in Mathematics in 1981 from the Rijksuniversiteit Leiden, and his Ph.D. in Mathematics in 1985 from the Free University in Amsterdam. He worked at Philips Electronics from 1985 till 1992. In 1990 he became Certified Fellow in Production and Inventory Management (CFPIM) of the American Production and Inventory Control Society (APICS). From 1990–1992 he was a teaching professor at Boston University Brussels. In 1991 he was appointed part-time professor in Industrial Mathematics at Technische Universiteit Eindhoven. Since 1992 he is a full time professor in Operations Management at the same university. He was Scientific Director of the Beta Research School for Operations Management and Logistics. In 2004 he was Edelman Award Finalist as member of the Philips Semiconductors team. His main research areas are Supply Chain Management and Concurrent Engineering with emphasis on quantitative analysis.

G. J. van Houtum is Professor of Maintenance, Reliability, and Quality at Eindhoven University of Technology since 2008. Prior to that he filled positions as assistant/associate professor at the University of Twente (1994–1998) and Eindhoven University of Technology (1999–2007). He obtained his M.Sc and Ph.D. degree in Applied Mathematics from Eindhoven University of Technology in 1990 and 1995, respectively. He is the scientific director of the Beta Research School for Operations Management and Logistics. His research is focused on: (1) Spare parts management; (2) Maintenance and availability management of capital goods; (3) The effect of design decisions on the total cost of ownership of capital goods. A significant part of this research is in collaboration with companies such as ASML, DAF, IBM, Nedtrain, Océ Technologies, Marel Stork, and Vanderlande Industries.