

# Controlling a divergent 2-echelon network with transshipments using the consistent appropriate share rationing policy

**Citation for published version (APA):**

Diks, E. B., & Kok, de, A. G. (1996). Controlling a divergent 2-echelon network with transshipments using the consistent appropriate share rationing policy. *International Journal of Production Economics*, 45(1-3), 369-379. [https://doi.org/10.1016/0925-5273\(95\)00007-0](https://doi.org/10.1016/0925-5273(95)00007-0)

**DOI:**

[10.1016/0925-5273\(95\)00007-0](https://doi.org/10.1016/0925-5273(95)00007-0)

**Document status and date:**

Published: 01/01/1996

**Document Version:**

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

**Please check the document version of this publication:**

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

[Link to publication](#)

**General rights**

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

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

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

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

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

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# Controlling a divergent 2-echelon network with transshipments using the consistent appropriate share rationing policy

E.B. Diks\*, A.G. de Kok

*Department of Mathematics and Computing Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, Netherlands*

---

## Abstract

Consider a two-echelon inventory system consisting of a central depot (CD) and a number of retailers. Only the retailers face customer demand. The CD is allowed to hold stock. In all stockpoints, the echelon inventory position is periodically raised to certain order-up-to-levels. At the central depot, incoming stock is allocated by using the *consistent appropriate share rationing (CAS)* policy. This means that this policy attempts to keep the ratio of the projected net inventory at any retailer over the systemwide projected net inventory constant at any time. The size of this ratio depends on the customer service level every retailer requires, and the behavior of the demand process.

When the orders arrive at the retailers, an instantaneous rebalancing of the total net stock of the retailers takes place, so as to maintain all end-stockpoint inventory at a balanced position. This rebalancing is realized by the transshipment of stock, assuming that the time to transship stock from one retailer to another is negligible compared to the replenishment lead time (lead time between CD and a retailer).

Object of this analysis is the determination of all the control parameters (integral order-up-to-level, parameters of allocation policy at the CD and of the rebalancing policy at the retailer), so as the desired (different) service levels are attained at the retailers at minimal expected total costs. Exact expressions are developed to determine these parameters. However we will use some heuristics to actually compute these parameters, because of this intractability of the exact expressions. For the parameters of the CAS policy at the CD we restrict the analysis to the case where the lead time from the CD to the retailers equals a review period. All analytical results are validated by Monte-Carlo simulation.

The model developed will be compared with the same model without periodic, instantaneous rebalancing at the retailer. This yields insight into the conditions under which transshipment could be useful.

**Keywords:** Inventory; Multi-echelon; Transshipment; Rationing

---

## 1. Introduction

So far a lot of research has been done to determine good (optimal) stocknorms for the control of

multi-echelon production and distribution networks. By 'optimal' we mean that the total costs (holding costs, distribution costs, etc.) are minimized under the condition that all the end-stockpoints (retailers) attain their pre-determined target service levels. In a lot of literature the stocknorms are determined by first defining a cost structure and

---

\* Corresponding author.

next finding cost-optimal policies. See for example [1–4]. The major disadvantage of this approach is that in order to guarantee the pre-determined target service levels, the penalty costs for shortages of stock have to be known. Unfortunately often these costs are unknown in practice and therefore the use of this approach is limited.

Another approach, which is introduced by De Kok [5] and Lagodimos [6], is a more 'service related' approach to determine the stocknorms. In De Kok [5] a planning procedure has been determined for a divergent two-echelon inventory model that operates according to a periodic review policy. No intermediate stocks are held at the central depot (CD), thus the CD serves merely as a coordinator. Later the model is extended in [7, 8] by allowing the CD to hold stock. In [9] some deficiencies are corrected in applying the logic proposed by De Kok [5].

A possible way to guarantee high service levels, but keeping low stocknorms is to allow lateral transshipments between the end-stockpoints. However we have to realize that by allowing these transshipments extra costs are involved:

- The information structure of the system probably has to be adapted, because the inventory position of all end-stockpoints have to be known at every review moment.
- Transshipping material from one end-stockpoint to another require extra distribution costs.

On the other hand when some end-stockpoints have excess inventory while others face shortages, lateral transshipment has gained in popularity as the appropriate recourse action for the avoidance of shortages. So the use of lateral transshipments depends on the trade-off between the extra costs involved and the ability to keep low stocknorms. Besides possible low stocknorms lateral transshipments also considerably reduce the *imbalance* in the inventory system. Imbalance can be seen as the deviation of the inventory position of retailers from the average inventory position of these retailers. Most models concerning divergent multi-echelon systems assume the impact of imbalance on the service levels to be negligible. Some literature analyze when this assumption is reasonable [5, 9].

In Tagaras [10] a two-echelon distribution system with two retailers employing an order-up-

to-level policy is considered. So called pooling (transshipping) between the retailers is allowed. The depot has infinite capacity and the replenishment lead time equals 0. Also the transshipment between retailers is assumed to be instantaneous. This model is characterized by complete pooling in that if there is an economic incentive to transship one item, then the maximum amount will be sent.

In this paper we consider a divergent two-echelon inventory system consisting of one CD (which is allowed to hold stock) serving  $N$  retailers. The CD uses a base-stock replenishment policy, i.e. every review period the CD orders enough from an outside supplier to bring the systemwide inventory position to a certain level. Upon receipt of this order, the CD allocates it to the retailers by using a *consistent appropriate share* (CAS) rationing policy. This policy attempts to keep the ratio of the projected net inventory at any retailer over the systemwide projected net inventory at any time equal to a pre-specified fraction. This model has already been analyzed by De Kok et al. [8]. In this paper we extend this model by allowing instantaneous transshipment of stock between retailers every review period. This model is only realistic when the lead times between retailers are negligible compared to the lead time from CD to the retailers. One of the main goals of this paper is to get insight for which instances the described transshipment model performs better compared to the model without these transshipments. This is done by comparing the results of a model without transshipments [4] with the model of this paper.

A similar study has been done by Jönsson and Silver [11]. They also compare a two-echelon system without transshipments with a model with transshipments (a redistribution system). They showed that a redistribution system becomes more advantageous in situations with high demand variability, a long planning horizon, many retailers, a high service level and short lead times.

The paper is organized as follows. In Section 2 we describe the considered system. In Section 3 we explain the rationing policy at the CD and the rebalancing policy. In Section 4 we indicate how to determine the control parameters of the rebalancing policy and the rationing policy at the CD. In Section 5 we present some numerical results and

compare them with the model of De Kok et al. [8]. Finally in Section 6 we give a few concluding remarks.

## 2. Model description

Consider an inventory distribution system consisting of a central depot (CD) and  $N$  retailers. Each retailer faces external demand, which is independent of the demand at the other retailers. When a retailer cannot satisfy customer demand, the shortage of on hand stock (physical stock) will be backordered. It is obvious that the retailers consider this inability to meet the demand undesirable, because of high penalty costs due to backordering. Therefore the retailer uses the following order policy so as to keep the amount backordered within bounds: At the end of every review period retailer  $n$  ( $n = 1, \dots, N$ ) places an order at the CD to bring the inventory position (stock on hand plus stock on order minus backorders) up to  $S_n$ . This order arrives after a positive, deterministic lead time of  $l$  review periods. Note that every retailer has the same replenishment lead time. After the arrival of a replenishment order a complete rebalancing of the net stock (stock on hand minus backorders) of all retailers takes place by instantaneous transshipment. The rebalancing policy we use in this paper corresponds to the CAS rationing policy of De Kok et al. [8]. In Section 3.2 the properties of this policy are elaborated. Immediately after rebalancing all retailers place an order at the CD to raise their inventory position to their order-up-to-level.

Besides the retailers, the CD also uses a periodic review ordering policy to replenish the stock at the CD, so as to meet the demands of the retailers. In this paper the duration of the review period at the CD and at the retailers are equal, and the review moments are synchronized. At the end of every review period the CD places an order to an outside supplier to bring the echelon inventory position (stock on hand at CD plus stock on order at CD plus inventory position of all retailers) to order-up-to-level  $S_0$ . This order arrives after a positive, deterministic lead time of  $L$  review periods. After receiving the order we can distinguish two possibilities:

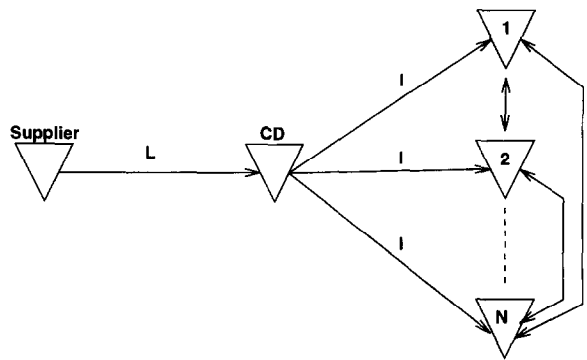


Fig. 1. Schematic representation of the inventory distribution system.

- The stock on hand at the CD is large enough to raise the inventory positions of the  $N$  retailers to their order-up-to-levels and the remainder is retained at the CD.
- The stock on hand at the CD is insufficient to meet the demand of all the retailers. When such a shortage occurs, we assume that the demand of the retailers which cannot be met is lost. An allocation rule will be used to ration the on hand stock over the retailers. Again a CAS rationing policy will be used, which will be thoroughly analyzed in Section 3.1. This material rationing at the CD takes place after the rebalancing at the retailers.

As we mentioned before the CD places an order at an outside supplier at the end of a review period. We assume that this supplier has an infinite capacity. Hence the echelon inventory position can always be raised to  $S_0$  and the CD never has to backorder.

The order-up-to-levels  $S_0$  and  $S_n$  have to be chosen in such a way that the disservice (the amount backordered per review period) of every retailer is acceptable. This is done by guaranteeing that retailer  $n$  gets a customer service level  $\beta_n^*$ . The service criterion considered in this paper is the fraction of demand satisfied directly from the stock on hand. This definition of the service level is widely used in practice [5, 6, 12].

Fig. 1 shows the inventory distribution system we analyze in this paper. The stockpoints are depicted by a triangle and the duration of the lead times

are depicted above the arrows. We assume that the demand retailer  $n$  faces during one review period has mean  $\mu_n$  and squared coefficient of variation  $c_n^2$ .

### 3. Analysis

Without loss of generality we assume that the duration of one review period corresponds to the duration of one period. In this section we use the following notation:

- $S_0$  := integral system order-up-to-level,
- $S_n$  := order-up-to-level of retailer  $n$ ,
- $X_t$  := stock on hand at CD at time  $t$  just after arrival replenishment order,
- $I_t^n$  := inventory position of retailer  $n$  at time  $t$  just before rationing,
- $\hat{I}_t^n$  := inventory position of retailer  $n$  at time  $t$  just after rationing,
- $J_t^n$  := net stock of retailer  $n$  at time  $t$  just before rebalancing,
- $\hat{J}_t^n$  := net stock of retailer  $n$  at time  $t$  just after rebalancing,
- $D_{t,t+k}$  := aggregate system demand in  $[t, t+k)$ ,
- $D_{t,t+k}^n$  := demand at retailer  $n$  in  $[t, t+k)$ ,
- $\mu_n$  := mean demand at retailer  $n$  during one period,
- $U_t^d$  := systemwide projected net inventory at the CD at time  $t+l+1$ ,
- $U_t^r$  := systemwide projected net inventory at the retailers at time  $t+1$ ,
- $p_n^d$  := allocation-fraction of retailer  $n$  of rationing at CD,
- $p_n^r$  := allocation-fraction of retailer  $n$  of rebalancing at retailers,
- $T$  := total expected stock transshipped between all the retailers every period,
- $T_n$  := expected stock transshipped by retailer  $n$  every period,
- $\Delta$  :=  $S_0 - \sum_{n=1}^N S_n$ ,
- $d_0$  :=  $\sum_{n=1}^N \mu_n$ ,
- $c$  :=  $S_0 - \Delta - d_0$ .

#### 3.1. Rationing policy at the CD

At the end of an arbitrary review period the CD raises the echelon inventory position to  $S_0$ . For

notational purposes we will refer to this point in time as  $t = 0$ . Because the depot lead time equals  $L$  periods, this order arrives at the end of period  $L$ . So the stock on hand at the CD after arrival of this order,

$$X_L = S_0 - D_{0,L} - \sum_{n=1}^N I_L^n. \quad (1)$$

If after the arrival of this order all the retailers want to raise their inventory position to their order-up-to-level, these has to hold:

$$X_L \geq \sum_{n=1}^N (S_n - I_L^n). \quad (2)$$

Using (1) and (2) we get

$$\Delta \geq D_{0,L}. \quad (3)$$

By choosing a  $\Delta \geq 0$  we can manipulate the role of the CD. When  $\Delta = 0$  the CD serves merely as a coordinator. This means that when the stock arrives at the CD it immediately is allocated to the retailers. While when  $\Delta = \infty$  the considered inventory system in fact reduces to  $N$  1-echelon systems working in parallel.

In order to explain the rationing policy properly we introduce  $U_t^d$ , which will be referred to as the *systemwide projected net inventory* at the end of period  $t+l+1$  just before a replenishment order arrives at the retailers (see [8]).  $U_t^d$  represents the best estimate for the sum of the projected net inventory of all retailers at the end of period  $t+l+1$  given the inventory position of the retailers at the end of period  $t$ . If condition (3) holds all the retailers can raise their inventory positions to their order-up-to-levels at time  $t=L$ . Hence the systemwide projected net inventory at  $t=L$  equals the planned cumulative safety stock of the retailers. In formula,

$$U_L^d = \sum_{n=1}^N (S_n - (l+1)\mu_n). \quad (4)$$

But because the CD has a finite capacity, condition (3) does not always hold. When this is the case the CD is unable to fulfil the demand of all the retailers, and therefore the systemwide projected net inventory does not coincide with the planned cumulative safety stock. Because the shortage in the CD equals

$D_{0,L} - \Delta$  the following equation holds:

$$U_L^d = \sum_{n=1}^N (S_n - (l+1)\mu_n) - (D_{0,L} - \Delta). \quad (5)$$

Combining (4) and (5) in one general formula yields,

$$U_L^d = S_0 - \Delta - (D_{0,L} - \Delta)^+ - (l+1)d_0, \quad (6)$$

where  $x^+ = \max\{0, x\}$ . Notice that because of this finite capacity of the CD the systemwide projected net inventory depends on the time.

The rationing policy which is used in this paper is a restricted version of the *appropriate share* (AS) rationing introduced by De Kok et al. [8]. This AS rationing can be viewed as an adaption of the allocation policy introduced by De Kok [5] for two-echelon depot-less networks. The purpose of AS rationing is to ensure that a pre-specified independent target service level can be attained at a retailer.

The policy we use in this paper is introduced in [8] as *consistent appropriate share* (CAS) rationing. This policy drastically reduces the number of decision variables involved, because it rations the depot inventory according to AS rationing but *chooses* the allocation fractions  $[p_n^d]$  such that

$$p_n^d = \frac{\hat{I}_L^n - (l+1)\mu_n}{\sum_{i=1}^N (\hat{I}_L^i - (l+1)\mu_i)}. \quad (7)$$

Clearly, we need that  $\sum_{n=1}^N p_n^d = 1$ . The rationale of this policy is that it attempts to keep the ratio of the projected net inventory at any retailer over the systemwide projected net inventory constant at any time.

Next we will derive an expression for the inventory position of retailer  $n$  after rationing,  $\hat{I}_L^n$ . As we mentioned before, if condition (3) holds all the retailers can raise their inventory position to their order-up-to-levels. Otherwise the on hand stock of the CD will be divided over the retailers. Hence,

$$\sum_{n=1}^N \hat{I}_L^n = \begin{cases} S_0 - D_{0,L}, & \Delta < D_{0,L} \\ \sum_{n=1}^N S_n & \Delta \geq D_{0,L}. \end{cases} \quad (8)$$

Again after some straightforward algebra, using (6)–(8), we obtain

$$\hat{I}_L^n = (l+1)\mu_n + p_n^d U_L^d. \quad (9)$$

This expression can be interpreted as follows: The inventory position of retailer  $n$  equals the expected demand at retailer  $n$  during the replenishment lead time plus a review period, plus a fraction of the systemwide projected net inventory (just before a rationing point in time).

If the depot inventory stock is rationed using the CAS rationing policy described above, this stock is not allocated consistently over the retailers. To illustrate this inconsistency consider a retailer with a large allocation-fraction. When  $U_L^d$  is positive this retailer profits because he gets a large part of the systemwide projected net inventory. However when  $U_L^d$  is negative this retailer is ‘punished’ because he gets a large part of the negative  $U_L^d$ . To deal with this problem we introduce an allocation-fraction  $q_n^d$  for negative  $U_L^d$ . From the above argument is clear that the following has to hold to provide a consistent rationing policy:

**Condition 3.1.** If for every  $n \in \{1, \dots, N\}$  we define  $q_n^d$  as a function of  $p_n^d$  then this function has to be monotonously decreasing in  $p_n^d$ .

Clearly also for  $q_n^d$  we require  $\sum_{n=1}^N q_n^d = 1$ . From Condition 3.1 it immediately follows that  $\hat{I}_L^n$  has to be adapted to ensure that an increasing  $p_n^d$  implies an increasing customer service level.

$$\hat{I}_L^n = (l+1)\mu_n + p_n^d (U_L^d)^+ - q_n^d (-U_L^d)^+. \quad (10)$$

In Section 4 we investigate the model for a particular form of  $q_n^d$ .

### 3.2. Rebalancing policy

In Section 3.1 we analyzed how goods coming from the supplier are allocated over the retailers. After this allocation at  $t = L$  the goods are shipped to the retailers. These orders arrive after  $l$  periods. During these periods retailer  $n$  faces a customer demand of  $D_{L,L+l}^n$  and the total net stock is rebalanced  $l - 1$  times. Hence the net stock of retailer  $n$  after the arrival of the order at

$t = L + l$  yields

$$J_{L+l}^n = \hat{I}_L^n - D_{L,L+l}^n + \sum_{t=L+1}^{L+l-1} (\hat{J}_t^n - J_t^n). \quad (11)$$

Notice that when  $l \geq 2$  the net stock  $J_{L+l}^n$  depends on the complete history of the system. However when the lead time from the CD to the retailers equals one review period ( $l = 1$ ),  $J_{L+l}^n$  only depends on  $\hat{I}_L^n$  and  $D_{L,L+l}^n$ .

Every time after the arrival of the orders at the retailers a complete rebalancing of the total net stock takes place by instantaneous transshipments. This is done by using a CAS rationing policy (see also Section 3.1). We know from the derivation of (6) that the inventory position of all the retailers together, after rationing at the CD, equals  $S_0 - \Delta - (D_{0,L} - \Delta)^+$ . After  $l$  periods the orders arrive at the retailers. During this period the total demand at the retailers equals  $D_{L,L+l}$ . So, the net stock of all the retailers, after arrival of the orders, equals  $S_0 - \Delta - (D_{0,L} - \Delta)^+ - D_{L,L+l}$ . Let us denote the systemwide projected net inventory at time  $t + 1$  by  $U_t^i$ , which represents the best estimate for the sum of the projected net inventory of all retailers at the end of period  $t + 1$  given the inventory position of the retailers at the end of period  $t$ .

$$U_{L+l}^i = S_0 - \Delta - (D_{0,L} - \Delta)^+ - D_{L,L+l} - d_0. \quad (12)$$

Now units are transshipped instantaneously in such a way that after rebalancing the net stock of retailer  $n$  yields,

$$\hat{J}_{L+l}^n = \mu_n + p_n^r (U_{L+l}^i)^+ - q_n^r (-U_{L+l}^i)^+. \quad (13)$$

This expression is similar to that of (10). The net inventory after rebalancing equals the expected demand retailer  $n$  has to face before a new order arrives, plus a fraction of  $U_{L+l}^i$ . Again we distinguish between a positive and a negative  $U_{L+l}^i$ . If  $U_{L+l}^i$  is positive, retailer  $n$  gets fraction  $p_n^r$ , otherwise he gets fraction  $q_n^r$ . Using an analogous argument like in the previous section we know condition 3.2 has to hold.

**Condition 3.2.** If for every  $n \in \{1, \dots, N\}$  we define  $q_n^r$  as a function of  $p_n^r$  then this function has to be monotonously decreasing in  $p_n^r$ .

By choosing  $p_n^r$  and  $q_n^r$  we are able to differentiate between the different retailers. The reason of favoring retailer  $n$  by choosing  $p_n^r$  relatively large, is based on the characteristics of retailer  $n$  (e.g. a large customer service level is demanded, or the customer demand is very unpredictable). How these differences between the retailers are expressed in the different sizes of the allocation-fractions is one of the main issues of this paper. In the next section is explained how to compute the allocation-fractions and in Section 5 we look at some examples.

It can easily be seen that the expected shortage at retailer  $n$  in the time-interval  $[L + l, L + l + 1)$  equals:

$$\mathbb{E}(D_{L+l,L+l+1}^n)^+ - \hat{J}_{L+l}^n - \mathbb{E}(-\hat{J}_{L+l}^n)^+. \quad (14)$$

Expression (14) represents the expected shortage at retailer  $n$  just before a new order arrives at  $t = L + l + 1$  minus the expected shortage at retailer  $n$  directly after rebalancing at  $t = L + l$ . Using the definition of the customer service level for retailer  $n$ ,  $\beta_n^*$ , yields

$$\beta_n^* = 1 - \frac{\mathbb{E}(D_{L+l,L+l+1}^n - \hat{J}_{L+l}^n)^+ - \mathbb{E}(-\hat{J}_{L+l}^n)^+}{\mu_n}, \quad n = 1, \dots, N. \quad (15)$$

#### 4. Determination of the control parameters

In this section we shortly elaborate on how the control parameters of the rebalancing policy and the rationing policy at the CD can be found. The control parameters  $S_0$ ,  $\{p_n^r\}$  and  $\{q_n^r\}$  are in fact determined by the service equations of (15). In [13] these service equations are rewritten to a more tractable relation between the known  $\beta_n^*$  and the unknown control parameters. This has been done by a two moment approximation of the exact expressions derived so far.

The service equations however consists of  $N$  equalities, while there are  $2N + 1$  unknown variables, namely  $\{p_n^r\}$ ,  $\{q_n^r\}$  and integral order-up-to-level  $S_0$ . To reduce this number of unknown variables, we choose  $q_n^r$  in the following straightforward way, such that condition (3.2) holds:

$$q_n^r = \frac{1 - p_n^r}{N - 1}. \quad (16)$$

Because there also holds  $\sum_{n=1}^N p_n^r = 1$ , the number of equations equals the number of unknown variables. Hence  $p_n^r$  is implicitly defined as a function of  $\beta_n^*$  and  $S_0$ .

Finally these control parameters  $\{p_n^r\}$  and  $S_0$  can be determined by using the heuristic algorithm of De Kok et al. [8]. This algorithm was introduced by De Kok [5]. Later some adaptations have been made by Verrijdt and De Kok [9]. When this algorithm halts we have good approximations of the integral order-up-to-level  $S_0$  and allocation-fractions  $\{p_n^r\}$  and  $\{q_n^r\}$ , such that every retailer gets a customer service level  $\beta_n^*$ .

Notice from (13) and (15) that the service level attained by retailer  $n$  is independent of  $p_n^d$ . This can be understood by considering a distribution system where the incoming stock is allocated poorly over the retailers with respect to the desired service levels of the retailers. Then the rebalancing policy reallocates a lot of stock every period to attain the desired service levels. Hence the disadvantage of a bad rationing policy at the CD is the large amount of stock which has to be transshipped every period. Therefore we shall determine  $\{p_n^d\}$  and  $\{q_n^d\}$  so as to minimize the total expected stock transshipped every period, which is denoted by  $T$ . This corresponds to minimizing the expected transshipment costs every period, when the costs of shipping stock from one retailer to another are equal. The size of  $T$  can also be interpreted as the total expected transshipments outflow of all the retailers. Hence,

$$T = \sum_{n=1}^N T_n, \tag{17}$$

where  $T_n$  equals the expected stock transshipped from retailer  $n$  to another every period. In formula,

$$T_n = \mathbb{E}(J_{L+1}^n - \hat{J}_{L+1}^n)^+. \tag{18}$$

In order to evaluate  $T_n$  we restrict the analysis to the case where  $l$  equals 1, because otherwise  $T_n$  is not tractable since  $J_{L+1}^n$  depends on the complete history of the system. Under this condition every retailer has exactly one outstanding order at the CD at any time. Now, using (6), and (10)–(13) we

can rewrite (18) as

$$\begin{aligned} T_n = & \mathbb{E}[\mu_n - D_{L,L+1}^n + p_n^d(c - (D_{0,L} - \Delta)^+ - d_0)^+ \\ & - q_n^d((D_{0,L} - \Delta)^+ + d_0 - c)^+ \\ & - p_n^r(c - (D_{0,L} - \Delta)^+ - D_{L,L+1})^+ \\ & + q_n^r((D_{0,L} - \Delta)^+ + D_{L,L+1} - c)^+]^+. \end{aligned}$$

$T_n$  consists of a lot of nested max-operators. This makes it (almost) impossible to derive an expression for  $T_n$  which is tractable. In order to keep the computations tractable we make the assumption  $p_n^d = q_n^d$  and  $p_n^r = q_n^r$ . Under this assumption Conditions 3.1 and 3.2 do not hold, and Monte-Carlo simulation reveals that the size of  $T$  differs significantly when we use this assumption. In spite of this the  $\{p_n^d\}$  for which  $T$  is minimized is well approximated by making this assumption. After adaptation of  $T_n$  using the property  $x^+ - (-x)^+ = x$  we get

$$T_n = \mathbb{E}(X_n + Y_n - Z_n - K_n)^+, \tag{19}$$

with

$$X_n = (p_n^r - p_n^d)(D_{0,L} - \Delta)^+,$$

$$Y_n = p_n^r \sum_{i \neq n} D_{L,L+1}^i,$$

$$Z_n = (1 - p_n^r)D_{L,L+1}^n,$$

$$K_n = (p_n^r - p_n^d)c + p_n^d d_0 - \mu_n.$$

So when the  $\{p_n^d\}$  are given we are able to compute the total expected stock  $T$  transshipped every period. But we are interested for which  $\{p_n^d\}$  this  $T$  is minimized. This problem corresponds to the so-called *resource allocation problem* [14]

$$\min T(p_1^d, \dots, p_N^d)$$

$$\text{s.t. } \sum_{n=1}^N p_n^d = 1,$$

$$0 \leq p_n^d \leq 1 \quad \text{for } n = 1, \dots, N.$$

Before we present an algorithm to solve this nonlinear optimization problem with  $N$  variables, we look at Theorem 4.1.



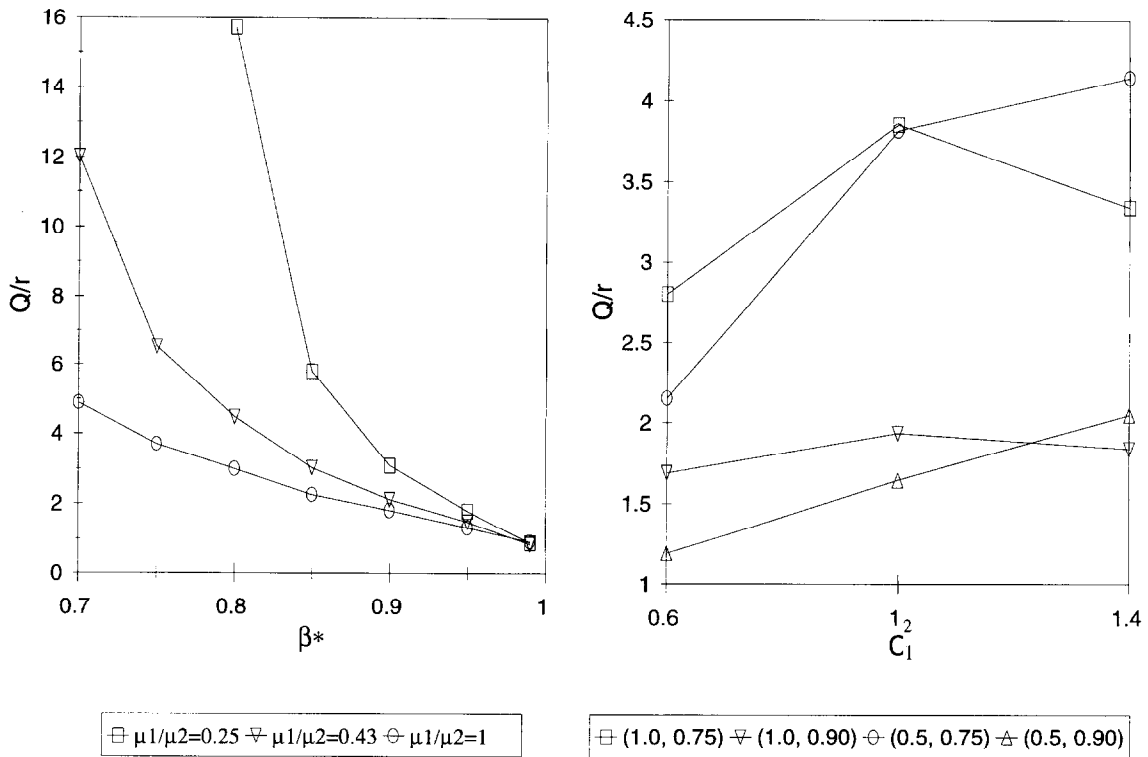


Fig. 2. A distribution system with  $N = 2, L = 6$  and  $l = 1$ . Also holds for (a)  $c_1^2 = c_2^2 = 0.6$ , and for (b)  $c_2^2 = 1.0$ .

**Theorem 4.1.** For all the retailers the expected stock transhipped every period is a strict convex function of the allocation-fraction  $p_n^d$ .

This theorem has been proven in [13]. This implies that our optimization problem is in fact a *convex resource allocation problem*. Using the convexity property of the objection-function we are able to solve the problem with a ‘conventional’ optimization algorithm. The gradient projection method of Rosen (see [15]) could for example be used.

Finally we have been able to determine all the control parameters of the two-echelon model. Unfortunately we have not been able to derive an analytic, tractable expression for  $T_n$ . Therefore we suggest to determine  $T_n$  by simulation, which is manageable because all the control parameters are known.

**5. Numerical results**

In this section we shall try to give insight as to when pooling between retailers could be profitable. In order to do this we have to know:

- The extra costs involved with the transshipments. Denote these costs per period by  $C_e$ .
- The (possible) reduction in costs by these transshipments. Denote this reduction in costs per period by  $C_r$ .

To simplify the computations we assume that  $C_e$  only consists of the distribution costs of shipping stock between retailers, and that these costs are equal for one retailer to every other. Hence,

$$C_e = \alpha T, \quad \alpha \geq 0,$$

where  $\alpha :=$  costs of shipping one unit of stock from one retailer to another per period.

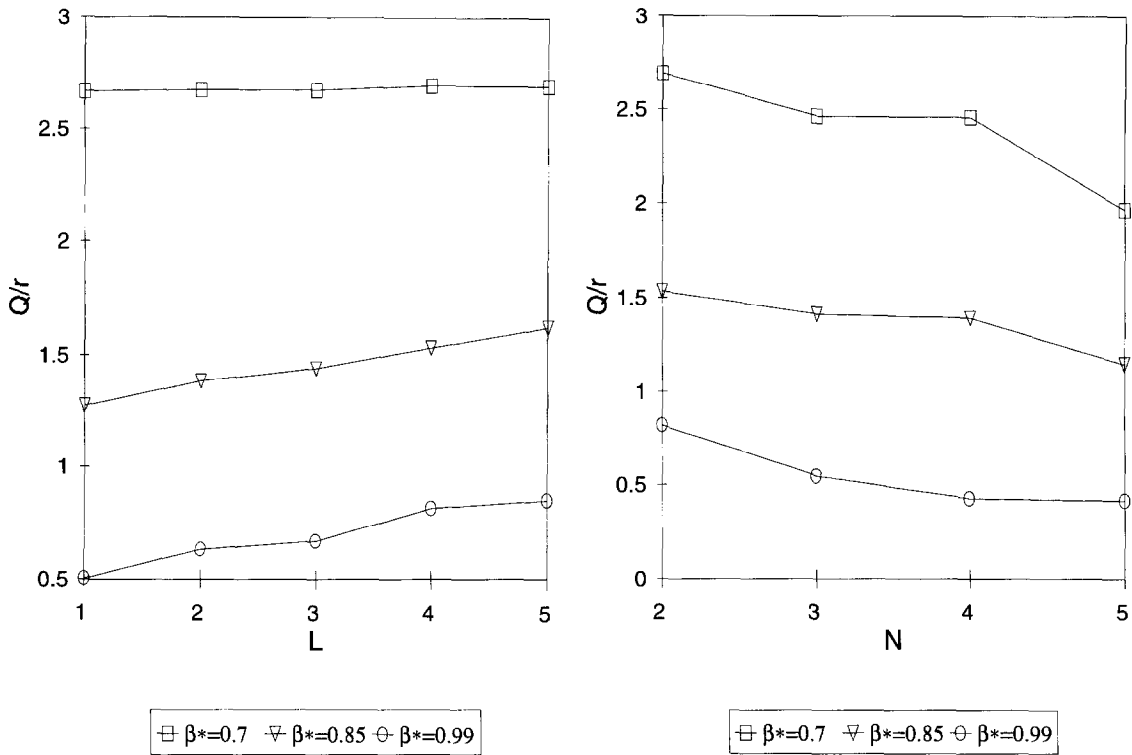


Fig. 3. (a) A distribution system with  $N = 2$ ,  $\mu_1/\mu_2 = 0.67$ ,  $c_1^2 = 0.6$ ,  $c_2^2 = 1.0$  and  $l = 1$ . (b) A distribution system with  $L = 4$  and  $l = 1$ .

Assume that the reduction in costs by transshipment is given by

$$C_r = h(S_0^{wt} - S_0), \quad h \geq 0,$$

where

$h$  := holding costs of one unit of stock per period.

$S_0^{wt}$  := integral system order-up-to-level for system without transshipment, which has been analyzed in [8].

Notice that now implicitly is assumed that the holding costs at the CD and at all the retailers are equal. For notational purposes we define  $r := \alpha/h$ , which is supposed to be a known constant. So transshipment becomes profitable when for the cost-coefficient  $Q := C_e/C_r$  satisfies

$$Q = r \frac{T}{S_0^{wt} - S_0} < 1.$$

The remainder of this section we devote to the relation between the system parameters and  $Q$ .

Furthermore we restrict ourselves to the case where every retailer demands the same customer service level  $\beta^*$  and  $\Delta = 0$ . Recall that  $\Delta = 0$  corresponds to a 2-echelon inventory system where the CD does not hold any stock, but merely serves as a coordinator. For cases where the CD holds stock and/or the retailers demand different customer service levels we refer to [13].

Fig. 2(a) depicts that  $Q$  decreases as the service level  $\beta^*$  increases, and that the more the mean demand per period of the retailers differs the larger  $Q$  gets when  $\beta^*$  is small.

Fig. 2(b) depicts the effect of the squared coefficient of variation of retailer 1,  $c_1^2$ , on  $Q$ . We vary  $c_1^2$  from 0.6 to 1.4 for the parameter set

$$\left\{ \left( \frac{\mu_1}{\mu_2}, \beta^* \right) \mid \frac{\mu_1}{\mu_2} \in \{0.5, 1.0\}, \beta^* \in \{0.75, 0.90\} \right\}.$$

Observe that when  $\mu_1/\mu_2 = 0.5$  pooling becomes less advantages with the increase of  $c_1^2$ . However

Table 1  
The characteristics of the demand processes of the participating retailers for the several distribution systems

Retailer $i$	$\mu_i$	$c_i^2$	$N = 2$	$N = 3$	$N = 4$	$N = 5$
1	5	1.0			X	X
2	10	0.6		X	X	X
3	10	0.6	X	X	X	X
4	15	1.0	X	X	X	X
5	20	1.4				X

when  $\mu_1/\mu_2 = 1$  the behavior of  $Q$  is completely different, because if also holds  $c_1^2 = 1.0$ , pooling is less advantageous than when  $c_1^2 = 1.4$ . This can be explained by the good performance of the model without transshipment when the retailers are completely identical.

In Fig. 3(a) we vary lead time  $L$  from 1 to 5, while keeping  $l$  constant. This figure shows that when the lead time between supplier and CD increases, pooling becomes less advantageous. This can be explained by the *statistical economies of scale* [16] since the model with transshipments profits from this effect in comparison with the model without transshipments. However when  $L$  increases this effect becomes relatively small.

Finally we like to analyze the effect of the number of retailers on  $Q$ . In Fig. 3(b) we try to get insight into this effect, which is complicated by the rapidly growing number of parameters. In Table 1 the characteristics of the demand processes of the retailers used in Fig. 3(b) are presented. A 'X' in the table means that this retailer is present in that distribution system. Fig. 3(b) shows that  $Q$  decreases with  $N$ . The extent of this decrease is determined by service level  $\beta^*$ . However also the demand characteristics of the added retailer influence the extent of this decrease. Notice that when  $N$  increases the probability of some retailers having excess inventory while others face shortages increase. Hence transshipments become more advantageous and  $Q$  decreases.

## 6. Conclusions

In this paper we considered a two-echelon distribution system consisting of a central

depot (CD) and a number of retailers. Every review period an instantaneous rebalancing of the total net stock takes place, by transshipping stock from one retailer to another. The rebalancing policy used in the CAS rationing policy of De Kok et al. [8], just like the rationing policy used at the CD. In this paper the integral system order-up-to-level and the allocation parameters of the rebalancing policy are determined, so that the desired (different) service levels are attained. Furthermore the parameters of the CAS rationing policy at the CD are determined such that the expected total costs are minimized. Unfortunately we have to restrict the analysis of these allocation-fractions at the CD to the case where the lead time from the CD to the retailers equals a review period, because otherwise the analysis becomes far more complicated.

A comparison of the described transshipment model with the same two-echelon distribution model without transshipment [8] shows that often the former model yields a considerably smaller integral system order-up-to-level. However, additional costs to rebalance the total net stock every review period are incurred. Therefore a condition under which transshipment could be useful is derived. This shows that the transshipment model becomes more advantageous in situations with many retailers, a high service level, mean demands per period of the same size, and the CD located as close as possible to the supplier. These results correspond to the results of Jönsson and Silver [11]. The effect of the demand variability is not known in advance. This strongly depends on all the system parameters.

Finally, a disadvantage of the model presented is that *every review period* a rebalancing between the retailers takes place. Therefore a model should be developed, which only transships when this is really necessary. This could be a topic for further research. Furthermore the model could be extended by taking different review periods for the CD and the retailers, or different lead times from the CD to the different retailers. Especially the latter extension will probably make the model very complicated.

## Acknowledgements

The authors wish to express their gratitude to two anonymous referees.

## References

- [1] Hoadley, H. and Heyman, D.P., 1977. A two-echelon inventory model with purchases dispositions shipments, returns and transshipments. *Naval. Res. Logist. Quart.*, 24: 1–19.
- [2] Karmarkar, U.S. and Patel, N.R., 1977. The one-period,  $n$ -location distribution problem. *Naval. Res. Logist. Quart.*, 24: 559–575.
- [3] Federgruen, A. and Zipkin, P., 1984. Approximations of dynamic, multilocation production and inventory problems. *Mgmt. Sci.*, 30: 69–84.
- [4] Federgruen, A., 1993. Centralized planning models for multi-echelon inventory systems under uncertainty, in: S.C. Graves et al. (Eds.), *Handbooks in OR and MS* Chapter 3, Vol. 4. Elsevier, Amsterdam, pp. 133–173.
- [5] de Kok, A.G., 1990. Hierarchical production planning for consumer goods. *Eur. J. Oper. Res.*, 45: 55–69.
- [6] Lagodimos, A.G., 1992. Multi-echelon service models for inventory systems under different rationing policies. *Int. J. Prod. Res.*, 30: 939–958.
- [7] Seidel, H.P. and de Kok, A.G., 1990. Analysis of stock allocation in a 2-echelon distribution system, Technical Report 098, CQM, Philips Electronics.
- [8] de Kok, A.G., Lagodimos, A.G. and Seidel, H.P., 1994. Stock allocation in a 2-echelon distribution network under service-constraints. *Int. J. Prod. Econom.*, submitted.
- [9] Verrijdt, J.H.C.M. and de Kok, A.G., 1996. Distribution planning for a divergent depotless two-echelon network under service constraints, *Eur. J. Oper. Res.*, 89: 341–354.
- [10] Tagaras, G., 1989. Effects of pooling on the optimization and service levels of two-location inventory systems. *IEEE Trans.*, 21: 250–257.
- [11] Jönsson, H. and Silver, E.A., 1987. Analysis of a two-echelon inventory control system with complete redistribution. *Mgmt. Sci.*, 33: 215–227.
- [12] Silver, E.A. and Peterson, R., 1985. *Decision Systems for inventory management and production planning*. Wiley, New York.
- [13] Diks, E.B. and de Kok, A.G., 1995. Transshipments in a divergent two-echelon network using the consistent appropriate share rationing policy, Technical report, Memorandum COSOR 95-31, Eindhoven University of Technology.
- [14] Ibaraki, T. and Katoh, N., 1988. *Resource Allocation Problems: Algorithmic Approaches*. The MIT Press, Cambridge, MA.
- [15] Bazaraa, M.S. and Shetty, C.M., 1979. *Nonlinear Programming-Theory and Algorithms*, Chapter 10. pp. 389–399.
- [16] Eppen, G. and Schrage, L., 1981. Centralized ordering policies in a multi-warehouse system with lead times and random demand. *Mgmt. Sci.*, 16: 51–67.
- [17] Donselaar K. van., 1990. Integral stock norms in divergent systems with lot-sizes. *Eur. J. Oper. Res.*, 45: 70–84.