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

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

Memorandum COSOR 94-40

Controlling a divergent 2-echelon network with transsshipments using the consistent appropriate share rationing policy (Working Paper)

E.B. Diks
A.G. de Kok

Eindhoven, December 1994
The Netherlands

Controlling a divergent 2-echelon network with transshipments using the consistent appropriate share rationing policy (Working Paper)

E.B. Diks and A.G. de Kok
Department of Mathematics and Computing Science
Eindhoven University of Technology

December 9, 1994

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 system projected net inventory constant at any time. The size of this ratio depends on the customer service level every retailer requires, and the behaviour 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 that 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 the intractability of the exact expressions. 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.

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. With '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 next finding cost-optimal policies. See for example Hoadley & Heyman [1977], Karmarkar & Patel [1977], Federgruen & Zipkin [1984] and Federgruen [1993]. 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, these costs are often unknown in practice and, therefore, the use of this approach is limited.

Another approach, which is introduced by De Kok [1990] and Lagodimos [1992], is a more 'service related' approach to determine the stocknorms. In De Kok [1990] 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. In Verrijdt & De Kok [1993] some deficiencies are corrected in applying the logic proposed by De Kok [1990]. Later the model is extended in Seidel & De Kok [1990] and De Kok, Lagodimos & Seidel [1994] by allowing the CD to hold stock.

Another way to guarantee high service levels, but keeping low stocknorms is to allow lateral transshipments between the end-stockpoints. However we have to realise 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 requires 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 with transshipments and the ability to keep low stocknorms and thereby low holding costs. 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-echelons systems assume the impact of imbalance on the service levels to be negligible. Some literature analyse when this assumption is reasonable [Donselaar, 1990; Verrijdt & De Kok, 1993].

In Tagaras [1989] 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 characterised by complete pooling in that if there is an economic incentive to transship one item, then the maximum amount will be sent. In our opinion the use of this model in practise is very limited, because of the very restrictive assumptions.

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 analysed by De Kok, Lagodimos & Seidel [1994]. 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 the model without transshipments (De Kok, Lagodimos & Seidel [1994]) with the model of this paper.

A similar study has been done by Jönsson & Silver [1987]. They also compare a two-echelon system without transshipments with a model with transshipments (a redistribution system). In their model the CD is not allowed to hold any stock and the considered review period is large compared to the lead times in the model. Furthermore every retailer has at most one outstanding order. However their model takes into account that transshipment of stock takes time. 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 present the rationing policy at the retailers and the rebalancing policy at the retailers. In Section 4 and Section 5 we determine the control parameters of the rebalancing policy respectively the rationing policy at the CD. In Section 6 we present some numerical results and compare some of the results with the model of Seidel & De Kok [1990] and De Kok, Lagodimos & Seidel [1994]. Finally in Section 7 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 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 as to keep the amount backordered within bounds: At the end of every review period retailer n ($n = 1, \ldots, N$) 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 (physical stock 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, Lagodimos & Seidel [1994]. In Section 3.2 the properties of this policy are elaborated. Immediately after rebalancing every retailer places an order at the CD to raise their inventory position to their order-up-to-level.

Besides the retailers also the CD uses a periodic review ordering policy to replenish the stock at the CD, in order 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 synchronised. At the end of every review period the CD places an order at an outside supplier to bring the echelon inventory (inventory position of CD and 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 distinct two possibilities:

- 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 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 analysed 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 order-up-to-level 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 β_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 practise [Silver & Peterson, 1985; Lagodimos, 1992; De Kok, 1990].

Figure 1 shows the inventory distribution system we analyse in this paper. The stockpoints are depicted by a triangle and the duration of the lead time from supplier to CD and the replenishment lead times are depicted above the arrows. We assume that the demand retailer n faces during one review period has mean μ_n and squared coefficient of variation c_n^2.

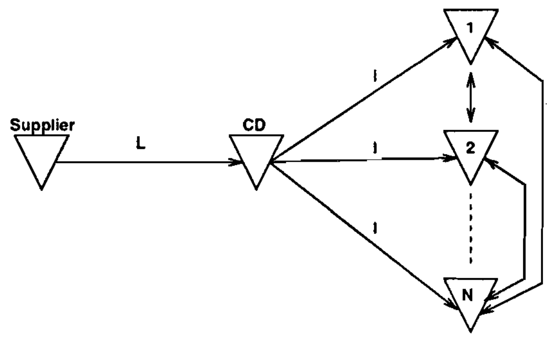

Figure 1: Schematic representation of the inventory distribution system.

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	:=	The integral order-up-to-level,
S_n	:=	The order-up-to-level of retailer n,
X_t	:=	Stock on-hand at CD at time t just after arrival order,
I_t^n	:=	The inventory position of retailer n at time t just **before** rationing,
$\hat{I}_t^n$	=	The inventory position of retailer n at time t just **after** rationing,
J_t^n	:=	The net stock of retailer n at time t just **before** rebalancing,
$\hat{J}_t^n$	=	The 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 $[t, t+k)$,
μ_n	:=	Mean demand at retailer n during one period,
U_t^d	:=	The systemwide projected net inventory at the CD at time $t+l+1$,
U_t^r	:=	The 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	:=	The total expected stock transshipped between all the retailers every period,
T_n	:=	The expected stock transshipped by retailer n every period,
Δ	:=	$S_0 - \sum_{n=1}^{N} S_n$,
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 common 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, equals

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

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

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

After rearranging (2) we get the following condition:

$$\Delta \geq D_{0,L} \,. \tag{3}$$

By choosing a $\Delta \geq 0$ we can manipulate with 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 parallelly.

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 De Kok, Lagodimos & Seidel [1994]). U_t^d represents the best estimate for the sum of the projected net inventory of all retailers at the end of period $t + l + 1$ as known 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,

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

But because the CD has a finite capacity, Condition (3) does not always hold. When this is the case the CD in 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,

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

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

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

The rationing policy which is used in this paper is a restricted version of the *Appropriate Share* (AS) rationing introduced by De Kok, Lagodimos & Seidel [1994]. This AS rationing can be viewed as an adaption of the allocation policy introduced by De Kok [1990] 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 De Kok, Lagodimos & Seidel [1994] as *Consistent Appropriate Share* (CAS) rationing. This policy drastically reduces the number of decision variables involved, because it rations the depot inventory such that

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

Clearly, we need that $\sum_{n=1}^{N} p_n^d = 1$. The rational 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,

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

Again after some straightforward algebra, using (6), (7) and (8), we obtain for $\hat{I}_L^n$:

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

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.

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. For every $n \in \{1, \ldots, N\}$ holds that if we define q_n^d as a function of p_n^d, this function has to be monotonously decreasing in p_n^d. □

Clearly also for q_n^d holds $\sum_{n=1}^{N} q_n^d = 1$. From Condition 3.1 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.

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

where $x^+ = max(0, x)$.

Implicitly we assume the imbalance assumption of De Kok [1990]. This means that for all t and $n \in \{1, \ldots, N\}$ holds $\hat{I}_t^n \geq I_t^n$. When at a certain time t this assumption is violated the allocation of stock to the different retailers must be adapted. Then the procedure described in De Kok [1990] can be followed.

3.2 Rebalancing policy

In Section 3.1 we analysed 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

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

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 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^r, which represents the best estimate for the sum of the projected net inventory of all retailers at the end of period $t+1$ as known at the end of period t. Thus U_{L+l}^r satisfies

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

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

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

This expression is similar to that of (9). 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}^r. Again we distinguish between a positive and a negative U_{L+l}^r. If U_{L+l}^r 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. For every $n \in \{1, \dots, N\}$ holds that if we define q_n^r as a function of p_n^r, 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 favouring retailer n by choosing p_n^d 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 6 we look at some examples.

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

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

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, β_n^*, yields

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

4 Determination of the control parameters of the rebalancing policy

In this section we elaborate on how the control parameters of the rebalancing policy can be determined. These control parameters S_0, $\{p_n^r\}$ and $\{q_n^r\}$ are implicitly defined by the service equations. However to determine these parameters we need a more tractable relation between the known β_n^* and the unknown control parameters. Therefore we fit a mixture of two Erlang distributions on the one-period demand distribution of every retailer. This means that the one-period demand of retailer n follows with probability α_1 an E_{k_1,λ_1} distribution and with probability $\alpha_2 := 1 - \alpha_1$ an E_{k_2,λ_2} distribution. In short we use the notation $ME(\alpha_1, \alpha_2, \lambda_1, \lambda_2, k_1, k_2)$. For more details of this fitting procedure we refer to Tijms [1994]. With F(.), f(.) we denote the cdf respectively the pdf of these mixed Erlang distributions.

To simplify the analysis of the service equations considerably we like to make a similar assumption as used in De Kok, Lagodimos & Seidel [1994].

Assumption 4.1. $X := (D_{0,L} - \Delta)^+ + D_{L,L+l} \sim ME(\beta_1, \beta_2, \phi_1, \phi_2, l_1, l_2)$. □

Notice that when Δ is small with regard to $\mathbf{E}D_{0,L}$ (e.g. $\Delta = 0$, which corresponds to a pure distribution system), $(D_{0,L} - \Delta)^+$ behaves almost like $D_{0,L}$. This implies that X behaves almost like the convolution of the random variables $D_{0,L}$ and $D_{L,L+l}$, which can be well approximated by a mixed Erlang distribution. When Δ is large with regard to $\mathbf{E}D_{0,L}$ (e.g. $\Delta = \infty$), $(D_{0,L} - \Delta)^+$ will be zero most of the time. This means that X behaves almost as $D_{L,L+l}$, which can also be well approximated by a mixed Erlang distribution. So only for Δ around $\mathbf{E}D_{0,L}$ the impact of this assumption is not predictable in advance. Monte-Carlo simulation has been performed to test how large the impact of this assumption is for several Δ. So far Monte-Carlo simulation reveals that indeed by making Assumption 4.1 only for Δ around $\mathbf{E}D_{0,L}$ the attained service levels differs a bit (at most 0.004) from the service levels which would be attained if we had not made this assumption.

Using (15) and Assumption 4.1 we are able to rewrite the service equations in terms of the system parameters. In Appendix A.1 an outline of this derivation is given. After considerable algebra we obtain,
if $q_n^r \neq 0$,

$$\mu_n(1-\beta_n^*) = \sum_{p=1}^{2}\alpha_p \sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-p_n^r)^j \psi_j^X(c,\lambda_p p_n^r)\tau_{i-j}\left(p_n^r,\lambda_p\right) + \sum_{p=1}^{2}\alpha_p \sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-q_n^r)^j \left[J_{i,j}(\lambda_p) - \psi_j^X(c,\lambda_p q_n^r)\tau_{i-j}(q_n^r,\lambda_p)\right], \tag{16}$$

if $q_n^r = 0$,

$$\mu_n(1-\beta_n^*) = \sum_{p=1}^{2}\alpha_p \sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-p_n^r)^j \psi_j^X(c,\lambda_p p_n^r)\tau_{i-j}(p_n^r,\lambda_p) + (1-F_{D_{n,1}}(\mu_n))(1-F_X(c)), \tag{17}$$

where definitions of $\psi_j^X(v,\eta)$, $\tau_j(x,\lambda)$ and $J_{i,j}(\lambda)$ are given in Appendix A.

Notice that the N above equalities contain $2N+1$ unknown variables, namely $\{p_n^r\}$, $\{q_n^r\}$ and 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:

$$q_n^r = \frac{1-p_n^r}{N-1}. \tag{18}$$

Because 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 β_n^* and S_0.

Now we present a heuristic algorithm which determines the integral order-up-to-level S_0 and the fractions p_n^r. This heuristic is proposed by De Kok [1990]. Later some adaptions have been made by Verrijdt & De Kok [1993], and De Kok, Lagodimos & Seidel [1994].

1. Initialization of S, e.g. $S := (L+l+1)d_0$.
2. Calculate $(p_n^r)^*$ for $n = 1,\ldots,N$, where $(p_n^r)^*$ is the value for which (16)–(17) holds with order-up-to-level S. This can be done by using bisection, since the customer service level β_n^* is an increasing function of $(p_n^r)^*$.
3. If $\sum_{n=1}^{N}(p_n^r)^* \simeq 1$ then stop, else
   if $\sum_{n=1}^{N}(p_n^r)^* > 1$ then increase S and go back to 2,
   if $\sum_{n=1}^{N}(p_n^r)^* < 1$ then decrease S and go back to 2.

When this heuristic algorithm has stopped we use S and $\{(p_n^r)^*\}$ as an approximation for the wanted integral order-up-to-level S_0 and fractions $\{p_n^r\}$, so that the service constraints are satisfied.

5 Determination of control parameters rationing policy

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 $\{q_n^d\}$ 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 outflow of all the retailers. Hence,

$$T = \sum_{n=1}^{N} T_n , \tag{19}$$

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

$$T_n = \mathbf{E}(J_{L+l}^n - \hat{J}_{L+l}^n)^+ . \tag{20}$$

In order to evaluate T_n we need a tractable expression for the net stock J_{L+l}^n. Unfortunately when $l > 1$, we know from (11) that J_t^n depends on the complete history of the system. Hence an analytic approach to determine $\{p_n^d\}$ becomes cumbersome. Therefore we shall distinguish between the case where $l = 1$ and $l > 1$.

5.1 Case where $l = 1$

In this case every retailer has exactly one outstanding order at the CD at any time. Because during the shipment of such an order no rebalancing takes place, the summation term of (11) equals 0. Now we can rewrite (20),

$$\begin{aligned} T_n = \; & \mathbf{E}\big[l\mu_n - D_{L,L+l}^n + p_n^d(c - (D_{0,L} - \Delta)^+ - ld_0)^+ - q_n^d((D_{0,L} - \Delta)^+ + ld_0 - c)^+ \\ & -p_n^r(c - (D_{0,L} - \Delta)^+ - D_{L,L+l})^+ + q_n^r((D_{0,L} - \Delta)^+ + D_{L,L+l} - c)^+\big]^+ . \end{aligned} \tag{21}$$

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. Therefore we suggest to assume $p_n^d = q_n^d,\ p_n^r = q_n^r$.

Under this assumption Condition 3.1 and 3.2 do not hold. Therefore Monte-Carlo simulation is used to get insight into the impact of this assumption on the $\{p_n^d\}$ for which T is minimized. This has been done for distribution systems with $N = 2$, $\Delta = 0$, $L = 6$, $l = 1$, $\mu_1/\mu_2 \in \{0.25, 1.0\}$, $c_1^2 = 1.0$, $c_2^2 \in \{0.6, 1.0, 1.4\}$ and $\beta_1^* = 0.9$, $\beta_2^* \in \{0.7, 0.8, 0.9, 0.99\}$. In 19 out of the 24 systems the optimal $\{p_n^d\}$ is not influenced by the assumption. The size of the minimal T, denoted by T_M, however differs significantly from T_A, the minimal T using the assumption. In Table 1 the relative difference (in percents) between T_M and T_A is depicted.

Because the optimal $\{p_n^d\}$ is hardly influenced by the assumption we will use this assumption in the remainder of this Section 5.1. Using the assumption and (21) yields

$$T_n = \mathbf{E}(X_n + Y_n - Z_n - K_n)^+ , \tag{22}$$

where
$$\begin{aligned} X_n &= (p_n^r - p_n^d)(D_{0,L} - \Delta)^+ , \\ Y_n &= p_n^r \sum_{i \neq n} D_{L,L+l}^i , \\ Z_n &= (1 - p_n^r) D_{L,L+l}^n , \\ K_n &= (p_n^r - p_n^d)c + l(p_n^d d_0 - \mu_n) . \end{aligned}$$

In Appendix A.2 we assume $|\, X_n \,|$, Y_n, Z_n to be distributed as a mixed Erlang distribution. This enables us to derive a tractable expression for T_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

		β_2^*			
μ_1/μ_2	c_2^2	0.7	0.8	0.9	0.99
0.25	0.6	14.1	9.4	0.7	3.3
	1.0	12.6	7.7	7.6	4.0
	1.4	12.3	7.2	14.4	4.2
1.00	0.6	2.1	2.5	1.0	0.7
	1.0	1.9	0.9	0.0	0.2
	1.4	4.1	2.1	0.5	0.0

Table 1: Relative difference between T_M and T_A for distribution systems with $\Delta = 0$, $N = 2$, $L = 6$, $l = 1$ $c_1^2 = 1.0$ and $\beta_1^* = 0.9$.

minimized. This problem corresponds to the so-called *resource allocation problem* [Ibaraki & Katoh, 1988],

$$\begin{aligned} & min\ T(p_1^d, \ldots, p_N^d) \\ \text{s.t.}\ & \sum_{n=1}^{N} p_n^d = 1, \\ & 0 \le p_n^d \le 1 \text{ for } n = 1, \ldots, N. \end{aligned}$$

Before we present an algorithm to solve this non-linear optimization problem with N variables, we look at Theorem 5.1.

Theorem 5.1. *For all the retailers hold that the expected stock transshipped every period is a strict convex function of the allocation-fraction* p_n^d. □

This theorem has been proven in Appendix B and implies that our optimization problem is in fact a *convex resource allocation problem.* Using the convexity property of the objective function we are able to solve the problem with a 'conventional' optimization algorithm. The gradient projection method of Rosen (see [Bazaraa & Shetty, 1979]) could for example be used.

5.2 Case where $l > 1$

In this case every retailer has several outstanding orders at the CD. During the time such an order is shipped from the CD to the retailer the total net stock at the retailers is rebalanced several times. This is expressed by the summation term in (11), which in general not equals zero. This term causes J_{L+l}^n to depend on the complete history of the system. Hence an anlytic approach becomes cumbersome. Therefore we suggest to use the following heuristic.

Heuristic 1: *All the allocation-fractions at the CD are equal to the allocation-fractions at the retailers.* □

The performance of this heuristic is analysed by simulation of several distribution systems. For every system we determine the $\{p_n^d\}$ which minimizes T. Because we have to estimate T for every set $\{p_n^d\}$ by simulation, we restrict ourselves to systems consisting of two retailers. Besides the minimal expected total stock transshipped every period, denoted by T_M, we estimate this minimal transshipment quantity when the heuristic is used, denoted by T_H. Now we are able to determine the relative difference δ (in percents) between T_M and T_H. The results are shown in the tables of Appendix C.

Table 3 depicts δ for distribution systems without any intermediate stock ($\Delta = 0$), $L + l = 7$ and $c_1^2 = 0.6$. This table indicates that the heuristic performs well (for $l > 1$ holds $\delta \leq 2\%$), especially when the target service level β^* is high (e.g. 0.99). Also notice the decrease of δ with μ_1/μ_2. The good performance of the heuristic is for a large part the result of the behaviour of T as a function of the $\{p_n^d\}$, because around the optimal $\{p_n^d\}$ the curve is flat.

Table 4 examines whether Δ influences the performance of the heuristic or not. This is done by analysing the same distribution systems like in Table 3, now however Δ equals 85% of the mean pipeline stock of the CD. For some systems δ is influenced by Δ, but the heuristic still performs well for all systems.

In Table 5 we investigate the impact of the retailers having different target service levels. Notice that the performance of the heuristic deteriorates when l decreases and the demand variability of the retailers increases. Hence the performance of this heuristic is only acceptable when l is large (e.g. $l = 6$) or/and the demand variability is low.

Table 6 examines the same distribution systems as in Table 5. Now however Δ equals 85% of the mean pipeline stock of the CD. A main difference between Table 5 and Table 6 is the very strong reduction in δ for those systems in Table 5 with very large δ (e.g. 22.8 and 9.7).

Unfortunately for some systems the performance of the heuristic is really poor. For those systems another heuristic or an analytical approach should be developed. This could be a topic for further research.

Finally we have accomplished 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.

6 Numerical results

In this section we shall try to give insight in 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 review period by C_e.
- The (possible) reduction in costs by these transshipments. Denote this reduction in costs per review 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 α := costs of shipping one unit of stock from one retailer to another.
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.
S_0^{wt} := integral order-up-to-level for system without transshipment, which has been analysed in De Kok, Lagodimos & Seidel [1994].

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-quotient $Q := C_e/C_r$ holds

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

First we shall analyse the relation between Q and the system parameters of distribution systems with retailers which all demand the same target service level. We restrict ourselves to this special class of distribution systems because when retailers demand significantly different service levels the imbalance assumption in the non-transshipment model causes a change in the attained service levels. Hence a comparison between the model with and without transshipment is impossible, due to the unknown actual S_0^{wt}. Furthermore we assume the CD not to hold any stock, i.e. $\Delta = 0$. The effect of holding stock at the CD is discussed afterwards.

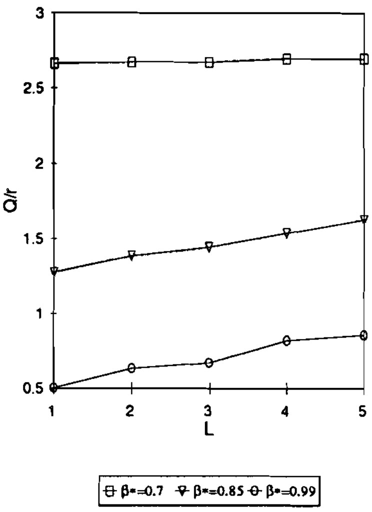

Figure 2: A distribution system with $N = 2$, $l = 1$, $\mu_1/\mu_2 = 0.67$, $c_1^2 = 0.6$ and $c_2^2 = 1.0$.

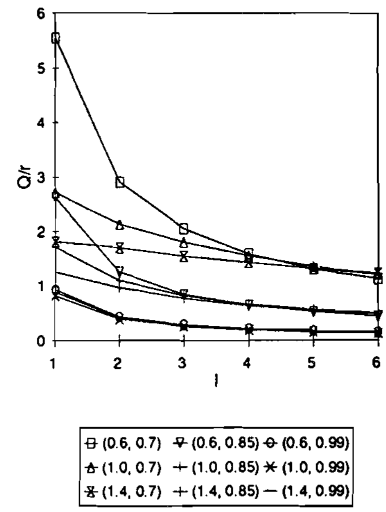

Figure 3: A distribution system with $N = 2$, $L + l = 7$, $\mu_1/\mu_2 = 0.67$ and $c_1^2 = 1.0$.

In Figure 2 we vary lead time L from 1 to 5, while keeping l constant. This figure shows that when the total lead time increases, pooling becomes less advantageous.

Figure 3 depicts the effect of the CD-location on Q, by varying lead time l and keeping $L + l$ constant. This has been done for the parameter set $\{(c_2^2, \beta^*) \mid c_2^2 \in \{0.6, 1.0, 1.4\}, \beta^* \in \{0.7, 0.85, 0.99\}\}$. Observe that pooling becomes more advantageous when the CD is positioned as close as possible near the supplier. Besides this also the demand variability (c_2^2) plays an important role when service level β^* and lead time l are not too large. The larger c_2^2 is the more advantageous pooling becomes. Most of the above observations can be explained by the effect *statistical economies of scale* [Eppen & Schrage, 1981]. At the end of an arbitrary review period the echelon inventory position is raised to the integral order-up-to-level. In the model of De Kok, Lagodimos & Seidel [1994] the allocation of this order to the individual retailers take place at the CD, after L periods. In the next l periods prior to the arrival at the retailers, every retailer has to overcome the fluctuations of the demand all by himself. In the transshipment model however the actual allocation to the individual retailers takes place after the arrival of the order at the retailers. That is why during the l periods prior to that arrival sudden fluctuations in the demand are overcome by all retailers.

Figure 4 depicts that Q decreases with the service level β^*, and that the more the mean demand per period of the retailers differs the larger Q gets when β^* is small.

Finally we like to analyse the effect of the number of retailers on Q. In Figure 5 we try to get insight into this effect, which is complicated by the rapidly growing number of parameters. In Table 2 the characteristics of the demand processes of the retailers used in Figure 5 are presented. A 'X' in the table means that this retailer is present in that distribution system. Figure 5 shows that Q

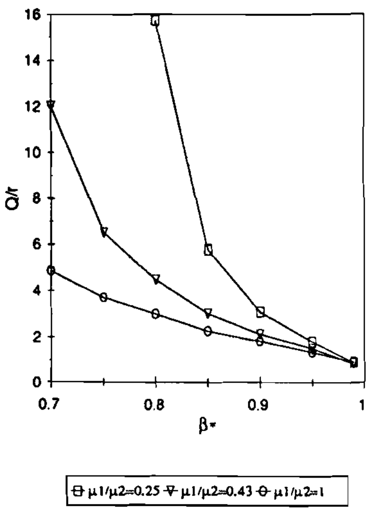

Figure 4: A distribution system with $N = 2$, $L = 6, l = 1$ and $c_1^2 = c_2^2 = 0.6$.

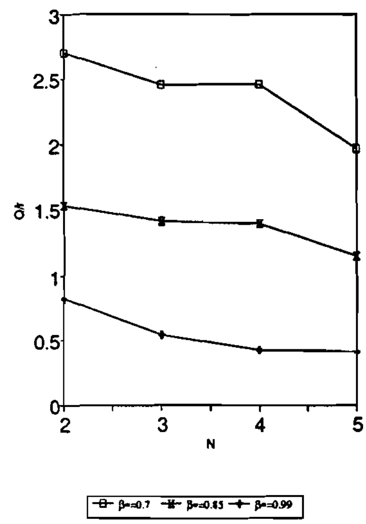

Figure 5: A distribution system with $L = 4$ and $l = 1$.

Retailer i	μ_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

Table 2: The characteristics of the demand processes of the retailers and the participating retailers for the several distribution systems.

decreases with N. The extend of this decrease is determined by service level β^*. However also the demand characteristics of the added retailer influence the extend 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.

So far all the distribution systems we considered did not hold any intermediate stock ($\Delta = 0$). Note that the stock at the CD at time t (after rationing) equals $(\Delta - D_{t-L,t})^+$. Hence when Δ increases, the mean amount of stock held at the CD increases. The stock held at the CD can not be used to satisfy customer demands, therefore the integral order-up-to-level S_0 has to increase to guarantee the service levels. On the other hand an increase of Δ diminishes the imbalance and simulation also reveals a decrease in T_H. Thus the optimal Δ depends on the trade off between the increase of S_0 (holding costs) and the decrease of T_M (transshipment costs). Figure 6 depicts this trade off for a distribution system with $N = 2$, $L = 3$, $l = 4$, $\mu_1 = 10$, $\mu_2 = 15$, $c_1^2 = 0.6$, $c_2^2 = 1.4$, $\beta_1^* = 0.95$ and $\beta_2^* = 0.8$.

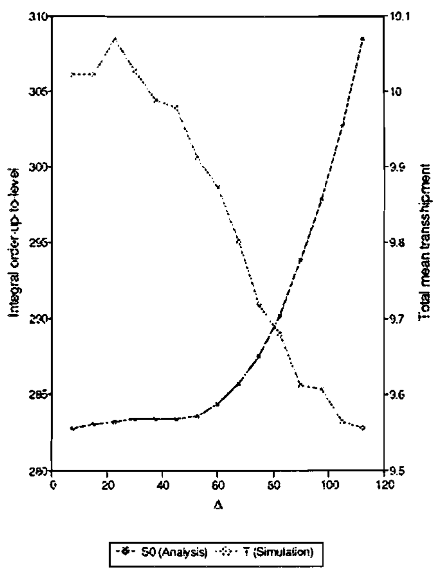

Figure 6: The trade off between S_0 and T_H for a distribution system with $N = 2$, $L = 3$, $l = 4$, $\mu_1 = 10$, $\mu_2 = 15$, $c_1^2 = 0.6$, $c_2^2 = 1.4$, $\beta_1^* = 0.95$ and $\beta_2^* = 0.8$.

For the optimal Δ holds

$$\frac{\partial S_0}{\partial \Delta} = -r\frac{\partial T_M}{\partial \Delta}.$$

More extensive research has to be done to get more insight in how the optimal Δ depends on the control parameters.

7 Conclusions

In this paper we considered a two-echelon distribution system consisting of a cental 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 is the CAS rationing policy of De Kok, Lagodimos & Seidel [1994], just like the rationing policy used at the CD. In this paper the integral order-up-to-level and the allocation parameters of the rebalancing policy are determined, so that the desired (different) service levels are attained. Furthermore we derived an analytical approach to determine the parameters of the CAS rationing policy at the CD, such that the mean total transshipment costs are minimized. This only holds when every retailer has at most one outstanding order at the CD all the time, i.e. $l = 1$, because if a retailer has more than one outstanding orders the analysis becomes cumbersome. For these cases we developed a heuristic which performs well for retailers with equal service levels.

A comparison of the described transshipment model with the same two-echelon distribution model without transshipment [De Kok, Lagodimos & Seidel, 1994] shows that often the former model yields a considerable smaller integral 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, a small total lead time and the CD located as close as possible to the supplier. These results correspond to the results of Jönsson & Silver [1987].

Finally, a disadvantage of the used model 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.

A Definitions

In this Appendix A we first give an outline on how equalities (16) and (17) can be derived. Secondly we introduce some assumptions to derive a tractable expression to compute T_n.
But before we do this we introduce the following notation:
If $X \sim ME(\alpha_1, \alpha_2, \lambda_1, \lambda_2, k_1, k_2)$ then holds,

$$C^X_{p,j}(x) = \alpha_p \frac{\lambda_p^{k_p}}{(\lambda_p + x)^{j+k_p}} \frac{(j + k_p - 1)!}{(k_p - 1)!},$$

$$\psi^X_j(v, \eta) = \sum_{p=1}^{2} \psi^X_{p,j}(v, \eta)\,, \quad \tilde{\psi}^X_{j,n}(v, \eta) = \psi^X_j(v, \eta) e^{\eta K_n}\,,$$

$$\psi^X_{p,j}(v, \eta) = \begin{cases} C^X_{p,j}(-\eta) \left[1 - \sum_{l=0}^{j+k_p-1} \frac{((\lambda_p - \eta)v)^l}{l!} e^{-(\lambda_p-\eta)v}\right] & \lambda_p \neq \eta \\ \alpha_p \frac{\lambda_p^{k_p}}{(k_p - 1)!} \frac{1}{j + k_p} v^{j+k_p} & \lambda_p = \eta \end{cases},$$

$$\theta^X_j(v, \phi) = \sum_{p=1}^{2} C^X_{p,j}(\phi) \sum_{l=0}^{j+k_p-1} \frac{((\lambda_p + \phi)v)^l}{l!} e^{-(\lambda_p+\phi)v}, \quad \bar{\theta}^X_{j,n}(v, \phi) = \theta^X_j(v, \phi) e^{-\phi K_n}\,.$$

Furthermore, we define

$$\tau_j(x,\lambda) = e^{-\lambda(\mu_n+xc)}\frac{j!}{\lambda^{j+1}}\sum_{k=0}^{j}\frac{(\lambda(\mu_n+xc))^k}{k!},$$

$$J_{i,j}(\lambda) = \sum_{p=1}^{2} J_{p,i,j}(\lambda),$$

$$J_{p,i,j}(\lambda) = \begin{cases} C^X_{p,j}(-\lambda q^r_n)\left[\tau_{i-j}(q^r_n,\lambda) - \sum_{l=0}^{j+l_p-1}\frac{\left(\frac{\phi_p}{q^r_n}-\lambda\right)^l}{l!}\tau_{i-j+l}\left(q^r_n,\frac{\phi_p}{q^r_n}\right)\right] & \phi_p \neq \lambda q^r_n \\ \beta_p\frac{\lambda^{l_p}}{(l_p-1)!}\frac{1}{j+l_p}\frac{\tau_{i+l_p}(q^r_n,\lambda)}{(q^r_n)^j} & \phi_p = \lambda q^r_n \end{cases},$$

$$J^{1,n}_{j,k}(v,\lambda) = \sum_{p=1}^{2} J^{1,n}_{p,j,k}(v,\lambda),$$

$$J^{1,n}_{p,j,k}(v,\lambda) = \begin{cases} C^{Y_n}_{p,j}(-\lambda)\left[\tilde{\theta}^{Z_n}_{k,n}(v,\lambda) - \sum_{l=0}^{j+l^n_p-1}\frac{(\phi^n_p-\lambda)^l}{l!}\sum_{s=0}^{l}\binom{l}{s}K_n^{l-s}\tilde{\theta}^{Z_n}_{k+s,n}(v,\phi^n_p)\right] & \phi^n_p \neq \lambda \\ \beta^n_p\frac{\phi_p^{l_p}}{(l_p-1)!}\frac{1}{j+l^n_p}\sum_{l=0}^{j+l^n_p}\binom{j+l^n_p}{l}K_n^{j+l^n_p-l}\tilde{\theta}^{Z_n}_{k+l,n}(v,\lambda) & \phi^n_p = \lambda \end{cases},$$

$$J^{2,n}_{j,k}(v,\phi) = \sum_{p=1}^{2} C^{Z_n}_{p,j}(\phi)\sum_{l=0}^{j+m^n_p-1}\frac{(-\eta^n_p-\phi)^l}{l!}\sum_{s=0}^{l}\binom{l}{s}K^{l-s}\tilde{\psi}^{X_n}_{k+s}(v,\eta^n_p).$$

A.1 Outline derivation service equations

In (15) the definition of the customer service level of retailer n, β^*_n, is given. Substitution of (13) in equality (15), yields

$$\beta^*_n = 1 - \frac{\mathbf{E}(Y_{1,n}-\mu_n)^+ - \mathbf{E}(Y_{2,n}-\mu_n)^+}{\mu_n}, \quad n = 1,\ldots,N. \tag{23}$$

with $Y_{1,n} = D^n_{L+l,L+l+1} - p^r_n(U^r_{L+l})^+ + q^r_n(-U^r_{L+l})^+$,
$Y_{2,n} = Y_{1,n} - D^n_{L+l,L+l+1}$.

Next we determine the expected shortage at the end of a period, $\mathbf{E}(Y_{1,n}-\mu_n)^+$. It can be shown that $\mathbf{E}(Y_{1,n}-\mu_n)^+$ yields

$$\begin{cases} \sum_{p=1}^{2}\alpha_p\sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-p^r_n)^j\psi^X_j(c,\lambda_p p^r_n)\tau_{i-j}(p^r_n,\lambda_p) + \\ \sum_{p=1}^{2}\alpha_p\sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-q^r_n)^j\left[J_{i,j}(\lambda_p) - \psi^X_j(c,\lambda_p q^r_n)\tau_{i-j}(q^r_n,\lambda_p)\right] + & q^r_n \neq 0 \\ \sum_{p=1}^{2}\beta_p\sum_{i=0}^{l_p-1}\frac{\left(\frac{\phi_p}{q^r_n}\right)^i}{i!}\tau_i\left(q^r_n,\frac{\phi_p}{q^r_n}\right), & \\ & \\ \sum_{p=1}^{2}\alpha_p\sum_{i=0}^{k_p-1}\frac{\lambda_p^i}{i!}\sum_{j=0}^{i}\binom{i}{j}(-p^r_n)^j\psi^X_j(c,\lambda_p p^r_n)\tau_{i-j}(p^r_n,\lambda_p) + & q^r_n = 0 \\ (1-F_{D^n_{L+l,L+l+1}}(\mu_n))(1-F_X(c)). & \end{cases} \tag{24}$$

The expected shortage immediately after rebalancing $\mathbf{E}(Y_{2,n}-\mu_n)^+$ yields

$$\begin{cases} \sum_{p=1}^{2} \beta_p \sum_{i=0}^{l_p-1} \frac{\left(\frac{\phi_p}{q_n^r}\right)^i}{i!} \tau_i\left(q_n^r, \frac{\phi_p}{q_n^r}\right) & q_n^r \neq 0 \\ 0 & q_n^r = 0 \end{cases} \tag{25}$$

Substitution of (24) and (25) in (23) yields the expressions (16)/(17).

A.2 Expression for T_n

In order to obtain a tractable expression for T_n we make the following assumption:

Assumption A.1. For the random variables of equation (22) holds

$$\begin{aligned} |X_n| &\sim ME(\alpha_1^n, \alpha_2^n, \lambda_1^n, \lambda_2^n, k_1^n, k_2^n), \\ Y_n &\sim ME(\beta_1^n, \beta_2^n, \phi_1^n, \phi_2^n, l_1^n, l_2^n), \\ Z_n &\sim ME(\gamma_1^n, \gamma_2^n, \eta_1^n, \eta_2^n, m_1^n, m_2^n). \end{aligned}$$

for all $n \in \{1, \ldots, N\}$. □

After considerable algebra we get,
if $p_n^r - p_n^d \geq 0$,

$$\begin{aligned} T_n = \ & 1_{\{K_n \leq 0\}}\left[(\mathbf{E}\mid X_n\mid + \mathbf{E}Y_n - K_n)F_{Z_n}(-K_n) - \mathbf{E}Z_n + \theta_1^{Z_n}(-K_n, 0)\right] + \\ & \sum_{p=1}^{2} \beta_p^n \sum_{i=0}^{l_p^n-1} \left(\mathbf{E}\mid X_n\mid + \frac{l_p^n - i}{\phi_p^n}\right) \frac{(\phi_p^n)^i}{i!} \sum_{j=0}^{i} \binom{i}{j} K_n^{i-j} \bar{\theta}_{j,n}^{Z_n}((-K_n)^+, \phi_p^n) + \\ & \sum_{p=1}^{2} \frac{\alpha_p^n}{\lambda_p^n} \sum_{i=0}^{k_p^n-1} (k_p^n - i) \frac{(\lambda_p^n)^i}{i!} \sum_{j=0}^{i} \binom{i}{j} (-1)^j \sum_{k=0}^{i-j} \binom{i-j}{k} K_n^{i-j-k} J_{j,k}^{1,n}((-K_n)^+, \lambda_p^n) , \end{aligned} \tag{26}$$

if $p_n^r - p_n^d < 0$,

$$\begin{aligned} T_n = \ & 1_{\{K_n \leq 0\}}\left[(\mathbf{E}Y_n - \mathbf{E}Z_n - K_n)F_{Z_n}(-K_n) - \mathbf{E}\mid X_n\mid + \theta_1^{|X_n|}(-K_n, 0)\right] + \\ & 1_{\{K_n \leq 0\}}\left[\sum_{p=1}^{2} \gamma_p^n \sum_{i=0}^{m_p^n-1} \left(-\mathbf{E}Y_n + \frac{m_p^n - i}{\eta_p^n}\right) \frac{(-\eta_p^n)^i}{i!} \sum_{j=0}^{i} \binom{i}{j} K_n^{i-j} \bar{\psi}_{j,n}^{|X_n|}(-K_n, \eta_p^n)\right] + \\ & 1_{\{K_n \leq 0\}}\left[\sum_{p=1}^{2} \frac{\beta_p^n}{\phi_p^n} \sum_{i=0}^{l_p^n-1} (l_p^n - i) \frac{(\phi_p^n)^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \sum_{k=0}^{i-j} \binom{i-j}{k} K_n^{i-j-k} J_{j,k}^{2,n}(-K_n, \phi_p^n)\right] + \\ & \sum_{p=1}^{2} \frac{\beta_p^n}{\phi_p^n} \sum_{i=0}^{l_p^n-1} (l_p^n - i) \frac{(\phi_p^n)^i}{i!} \sum_{j=0}^{i} \binom{i}{j} \theta_j^{Z_n}(0, \phi_p^n) \sum_{k=0}^{i-j} \binom{i-j}{k} K_n^{i-j-k} \bar{\theta}_{k,n}^{|X_n|}((-K_n)^+, \phi_p^n) . \end{aligned} \tag{27}$$

B Proof of Theorem 5.1

In this section we give the proof of Theorem 5.1. For that purpose we determine the second derivative of T_n to p_n^d, where

$$T_n = \mathbf{E}(X_n + Y_n - Z_n - K_n)^+ . \tag{28}$$

After conditioning of Y_n and Z_n in (28) we obtain

$$T_n = \int_0^\infty \int_0^\infty \mathbf{E}\left[(p_n^r - p_n^d)(D_{0,L} - \Delta)^+ - v(y, z, p_n^d)\right]^+ \, dF_{Y_n}(y) dF_{Z_n}(z) , \tag{29}$$

where $v(y, z, p_n^d) = z - y + l(p_n^d d_0 - \mu_n) + (p_n^r - p_n^d)c$.
Taking the second derivative of T_n to p_n^d yields

$$\frac{\partial^2 T_n}{\partial (p_n^d)^2} = \int_0^\infty \int_0^\infty H(y, z, p_n^d) \, dF_{Y_n}(y) dF_{Z_n}(z) , \tag{30}$$

where $H(y,z,p_n^d) = \dfrac{\partial^2 \mathbf{E}\left[(p_n^r - p_n^d)(D_{0,L}-\Delta)^+ - v(y,z,p_n^d)\right]^+}{\partial (p_n^d)^2}$.

By conditioning on $(D_{0,L}-\Delta)^+$ and case distinction of $p_n^r - p_n^d$ we are able to get an explicit expression for $H(y,z,p_n^d)$. After some straightforward algebra, we get

$$H(y,z,p_n^d) = \begin{cases} (-v(y,z,p_n^d))^+ & p_n^r = p_n^d \\ -1_{\{v(y,z,p_n^d)<0\}} \dfrac{(z-y+l(p_n^r d_0 - \mu_n))^2}{(p_n^r - p_n^d)^3} f_{(D_{0,L}-\Delta)^+}\left(\dfrac{v(y,z,p_n^d)}{p_n^r - p_n^d}\right) & p_n^r < p_n^d \\ 1_{\{v(y,z,p_n^d)>0\}} \dfrac{(z-y+l(p_n^r d_0 - \mu_n))^2}{(p_n^r - p_n^d)^3} f_{(D_{0,L}-\Delta)^+}\left(\dfrac{v(y,z,p_n^d)}{p_n^r - p_n^d}\right) & p_n^r > p_n^d \end{cases} . \tag{31}$$

Define,

$G_1 := \{(y,z) \in \mathbf{R}_+^2 \mid v(y,z,p_n^d) < 0\}$,

$G_2 := \{(y,z) \in G_1 \mid z - y + l(p_n^r d_0 - \mu_n) \neq 0\}$,

$G_3 := \{(y,z) \in \mathbf{R}_+^2 \mid v(y,z,p_n^d) > 0, z - y + l(p_n^r d_0 - \mu_n) \neq 0\}$.

It is easy to see that G_1, G_2 and G_3 are non-empty sets. Using equations (30) and (31) yields

$$\frac{\partial^2 T_n}{\partial (p_n^d)^2} = \iint_G H(y,z,p_n^d)\, dF_{Y_n}(y) dF_{Z_n}(z) ,$$

where G equals G_1, G_2 or G_3, depending on p_n^d and p_n^r.

Therefore G is non-empty, hence there exist a $(y,z) \in G$. For this (y,z) holds that $H(y,z,p_n^d) > 0$. Hence the second derivative of T_n to p_n^d is positive. So T_n is a strict convex function of p_n^d. Then it is clear that the expected stock transshipped every period, T, is a convex function of the allocation-fraction p_n^d, because the sum of strict convex functions is also strict convex. □

C Tables on testing performance Heuristic 1

			$l=1$				$l=3$				$l=6$			
μ_1/μ_2	c_2^2	β^*	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99
	0.6		1.1	1.2	1.1	0.8	0.4	1.2	1.7	1.0	0.0	0.3	1.1	0.4
0.25	1.0		2.2	2.0	1.5	0.0	1.9	1.8	0.8	0.0	1.4	1.5	0.4	0.0
	1.4		2.4	1.8	0.9	0.0	1.4	0.5	0.0	0.0	0.7	0.3	0.1	0.0
	0.6		0.4	0.4	0.3	0.0	0.0	0.1	0.3	0.0	0.2	0.0	0.4	0.0
0.67	1.0		0.6	0.2	0.0	0.0	0.5	0.4	0.3	0.0	1.0	1.2	0.3	0.0
	1.4		0.3	0.0	0.0	0.1	0.3	0.2	0.0	0.4	0.8	0.3	0.1	0.1
	0.6		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.00	1.0		0.0	0.0	0.1	0.2	0.4	0.2	0.1	0.2	0.7	0.6	0.2	0.0
	1.4		0.0	0.1	0.3	0.4	0.2	0.0	0.0	0.8	0.7	0.4	0.1	0.2

Table 3: Relative difference δ for distribution systems with $\Delta = 0$, $N = 2$, $L + l = 7$ and $c_1^2 = 0.6$.

			$l=1$				$l=3$				$l=6$			
μ_1/μ_2	c_2^2	β^*	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99
	0.6		1.3	1.3	1.2	0.8	0.6	1.4	1.7	0.9	0.0	0.2	0.7	0.1
0.25	1.0		2.4	1.9	1.3	0.0	1.7	1.4	0.4	0.0	0.8	1.0	0.1	0.0
	1.4		2.3	1.3	0.4	0.0	0.8	0.1	0.1	0.0	0.3	0.1	0.0	0.0
	0.6		0.6	0.3	0.2	0.0	0.0	0.1	0.3	0.0	0.2	0.0	0.2	0.0
0.67	1.0		0.6	0.2	0.0	0.0	0.6	0.6	0.3	0.0	0.4	0.8	0.4	0.0
	1.4		0.2	0.0	0.0	0.1	0.5	0.2	0.0	0.3	0.8	0.8	0.3	0.0
	0.6		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1.00	1.0		0.0	0.0	0.1	0.2	0.6	0.5	0.1	0.2	1.2	0.9	0.6	0.0
	1.4		0.0	0.1	0.3	0.5	0.4	0.1	0.0	0.7	1.3	1.1	0.3	0.1

Table 4: Relative difference δ for distribution systems with $\Delta = 0.85Ld_0$, $N = 2$, $L + l = 7$ and $c_1^2 = 0.6$.

			$l=1$				$l=3$				$l=6$			
μ_1/μ_2	c_2^2	β_2^*	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99
	0.6		0.8	1.1	1.0	4.2	0.5	0.9	1.7	0.7	0.0	0.1	0.7	0.1
0.25	1.0		2.2	2.7	1.6	3.7	1.8	2.3	1.7	2.3	0.5	0.7	1.1	0.4
	1.4		3.4	4.1	1.1	22.8	2.8	3.7	0.3	9.7	1.1	1.5	0.1	4.0
	0.6		1.2	1.2	0.0	4.2	0.5	0.3	0.0	2.8	0.0	0.0	0.2	0.9
1.00	1.0		3.0	2.6	0.0	3.8	1.2	0.8	0.0	3.8	0.0	0.0	0.0	1.6
	1.4		4.7	3.7	0.2	3.8	1.8	1.0	0.0	4.6	0.1	0.0	0.1	2.2

Table 5: Relative difference δ for distribution systems with $\Delta = 0$, $N = 2$, $L + l = 7$, $c_1^2 = 1.0$ and $\beta_1^* = 0.9$.

			$l=1$				$l=3$				$l=6$			
μ_1/μ_2	c_2^2	β_2^*	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99	0.7	0.8	0.9	0.99
	0.6		0.9	1.3	1.1	3.9	0.6	1.1	1.8	0.8	0.0	0.1	0.4	0.1
0.25	1.0		2.3	3.0	1.4	4.2	2.0	2.6	1.3	1.9	0.3	0.4	0.6	0.3
	1.4		3.9	4.8	0.7	3.9	3.3	4.0	0.1	2.9	0.8	1.0	0.0	0.4
	0.6		1.4	1.3	0.1	4.5	0.3	0.2	0.1	2.5	0.2	0.3	0.6	0.3
1.00	1.0		3.2	2.7	0.0	4.2	0.9	0.5	0.0	3.3	0.0	0.1	0.0	0.7
	1.4		5.2	3.8	0.2	4.1	1.7	0.6	0.0	4.2	0.0	0.0	0.3	1.0

Table 6: Relative difference δ for distribution systems with $\Delta = 0.85Ld_0$, $N = 2$, $L + l = 7$, $c_1^2 = 1.0$ and $\beta_1^* = 0.9$.

References

BAZARAA, M.S., AND C.M. SHETTY [1979], *Nonlineair Programming-Theory and Algorithms*, Chapter 10, 389–399.

DONSELAAR, K. VAN [1990], Integral stock norms in divergent systems with lot-sizes, *European journal of operations research* **45**, 70–84.

EPPEN, G., AND L. SCHRAGE [1981], Centralized ordering policies in a multi-warehouse system with lead times and random demand, *Managment Sciences* **16**, 51–67.

FEDERGRUEN, A. [1993], Centralized planning models for multi-echelon inventory systems under uncertainty, in: S.C. Graves et al. (ed.), *Handbooks in OR and MS*, Elsevier Science Publishers, Chapter 3, 133–173, Vol. 4.

FEDERGRUEN, A., AND P. ZIPKIN [1984], Approximations of dynamic, multilocation production and inventory problems, *Management Science* **30**, 69–84.

HOADLEY, H., AND D.P. HEYMAN [1977], A two-echelon inventory model with purchases dispositions shipments, returns and transshipments, *Naval Res. Logist. Quart.* **24**, 1–19.

IBARAKI, T., AND N. KATOH [1988], *Resource Allocation Problems: Algoritmic Approaches*, The MIT Press.

JÖNSSON, H., AND E.A. SILVER [1987], Analysis of a two-echelon inventory control system with complete redistribution, *Management Science* **33**, 215–227.

KARMARKAR, U.S., AND N.R. PATEL [1977], The one-period, n-location distribution problem, *Naval. Res. Logist. Quart.* **24**, 559–575.

KOK, A.G. DE [1990], Hierarchical production planning for consumer goods, *European Journal of Operational Research* **45**, 55–69.

KOK, A.G. DE, A.G. LAGODIMOS, AND H.P. SEIDEL [1994], *Stock allocation in a 2-echelon distribution network under service-constraints*, Technical report, TUE/BDK/LBS/94-03.

LAGODIMOS, A.G. [1992], Multi-echelon service models for inventory systems under different rationing policies, *International Journal of Production Research* **30**, 939–958.

SEIDEL, H.P., AND A.G. DE KOK [1990], *Analysis of stock allocation in a 2-echelon distribution system*, Technical Report 098, CQM, Philips Electronics.

SILVER, E.A., AND R. PETERSON [1985], *Decision Systems for inventory management and production planning*, John Wiley and Sons, New York.

TAGARAS, G. [1989], Effects of pooling on the optimization and service levels of two-location inventory systems, *IEEE Transactions* **21**, 250–257.

TIJMS, H.C. [1994], *Stochastic Models: an Algoritmic Approach*, Wiley, Chichester.

VERRIJDT, J.H.C.M., AND A.G. DE KOK [1993], *Distribution Planning for a Divergent 2-Echelon Network without Intermediate Stocks under Service Restrictions*, Technical report, TUE/BDK/LBS/93-27.