On the use of break quantities in multi–echelon distribution systems

Rommert Dekker, Hans Frenk, Marcel Kleijn, Nanda Piersma

Erasmus University Rotterdam, The Netherlands

Ton de Kok

Eindhoven University of Technology, The Netherlands

October 1995

Abstract

In multi–echelon distribution systems it is usually assumed that demand is only satisfied from the lowest echelon. In this paper we will consider the case where demand can be satisfied from any level in the system. However, then the problem arises of how to allocate orders from customers to the different locations. A possible way of dealing with this problem consists of using a so–called break quantity rule. This easy implementable rule is to deliver every order with a size exceeding the break quantity from a higher echelon. The use of the break quantity rule now results in a reduction of the demand variability at the retailer and hence less safety stocks need to be held. The concept is studied for a two–echelon distribution system, consisting of one warehouse and one retailer, where the inventory at the retailer is controlled by an order up to level policy, and where at the warehouse there is enough inventory to satisfy all orders from the retailer and the customers. For this system an approximation for the long run average costs as a function of the break quantity is derived, and an algorithm is presented to determine the cost–optimal break quantity. Computational results indicate that the break quantity rule can lead to significant cost reductions.

Keywords: Break quantity rule, inventory, multi–echelon distribution systems.

1 Introduction

In multi–echelon distribution systems, management is primarily concerned with meeting customer demand in a timely way at minimum cost. In practice, customer demand is often not known in advance, and can exhibit large fluctuations over time. Such fluctuations typically occur when the demand process consists of many customers ordering small batches and few customers ordering large batches. This may lead to high safety stock levels with large associated costs.

A way to reduce the variability of the demand is by handling large orders in a special way. We introduce a so–called *break quantity* to distinguish between small and large orders. In many cases customers ordering large batches do not have an immediate need for delivery, since their orders are used as a replenishment for an inventory system, or their order sizes are larger than necessary because of quantity discounts. Two management procedures, in which large orders are handled separately, are considered in De Kok [13] and Nass, Dekker and Van Sonderen–Huisman [15]. The first procedure is to split the large orders into smaller batches, which are subsequently delivered. This way the variability of the order sizes is reduced, and part of future demand will be known beforehand, which will enable a reduction of the safety stock levels. The second procedure, which we focus on in this paper, is to supply the large orders, at a later stage, from a higher level in the distribution system. In this paper it will be shown that this tactical rule leads to a reduction of the safety stocks at the lowest level of the distribution system, thus decreasing the long run average costs. The main reasons for this are the fact that the break quantity filters out the peaks in the demand, and the fact that the total demand for the lowest level decreases. Another advantage is the reduction of handling and transportation costs, due to the possibility of delivering large orders directly to the customers, thereby skipping one or more nodes in the distribution chain. The effect of direct shipping strategies on the transportation costs has been studied extensively in literature (e.g. Burns *et al.* [4], Gallego and Simchi–Levi [9], Sussams [19]). However, the effect on the inventory costs has only been analyzed for deterministic demand. A last advantage of the break quantity concept is its simplicity: it can easily be implemented in practice, especially when only statistical information on future demand is known, and upon order entry immediate feedback on the supply source can be given. However, besides advantages of handling large orders separately there are also disadvantages. A negative aspect of the break quantity rule is the (possible) occurrence of a setup cost for every large order. Another disadvantage is that the response time to the customer increases, since in general the distance from the lowest level of the distribution system to the customer is smaller than from a higher level (e.g. Axsäter [3]). An increase in the response time to the customers with large orders will induce penalty costs, that may represent price rebates to persuade customers to accept the longer delivery time. The idea of using price rebates to compensate longer delivery times was already presented by Jaikumar and Kasturi Rangan [10, 12]. They analyzed the question of how to set price rebates that allocate orders to different stages in a multi–echelon distribution system. The rebates differ from one echelon to the other and they depend on the delivery location, so customers are forced to choose the location from which they take possession of a product. The main goal of their papers was to structure the optimal buying arrangement such that customers minimize their costs simultaneously, while the manufacturer maximizes profit. They did not analyze the effect on the inventories, and they only considered the case where all orders were known beforehand.

The analysis in this paper is motivated by a case study for a company in Western–Europe that delivers self–adhesive materials to customers through a multi–echelon distribution system (Nass, Dekker and Van Sonderen–Huisman [15]). The multi–echelon distribution system is composed of Plant Service Centres (PSC's), which are associated to production plants and thus have both production and warehouse facilities, and of Distribution Centres (DC's), which are only used as stockpoints. In this system the customers, from all over Europe, could be served either by the nearest PSC or by the nearest DC. To operate this system the company needed a tactical rule to allocate orders from customers to different locations. The company used *the break quantity rule*, which was implemented such that orders with a size smaller than or equal to the break quantity are delivered from the nearest DC within 48 hours, and orders with a size exceeding the break quantity are delivered from the nearest PSC within 10 days. Using this concept it was possible to reduce the average system costs considerably.

In this paper we will study the effects of the break quantity rule for a two–echelon distribution system consisting of one warehouse and one retailer. The main goal of this paper is to show the effect of the break quantity rule on the inventory costs at the retailer. Therefore we make the simplifying assumption that the warehouse can always satisfy replenishment orders from the retailer and large orders from customers. Hence, we do not take into account the effect of the break quantity rule on the inventory at the warehouse. However, it is assumed that additional costs are incurred for delivering a large order separately, which may include price rebates and setup costs. In the next section an approximation for the minimum long run average system costs as a function of the break quantity are derived. Moreover, an easy and tractable condition under which the break quantity rule reduces the total long run average costs is given. In Section 3 an algorithm that determines the optimal break quantity is presented, while it is assumed that the management puts a restriction on the minimum number of customers served by the retailer. Finally, in Section 4 computational results are reported, to illustrate the influence of the break quantity rule on the average costs for a wide range of parameter settings. The results indicate that in many of the situations analyzed, a significant reduction in the long run average costs can be obtained. Moreover, it is shown that using the approximated cost function leads to similar results to those obtained by exact calculation.

2 The model

Consider a single product, two–echelon distribution system, consisting of one warehouse and one retailer. Assume that at the retailer the inventory is reviewed at discrete time points and that every review an order is placed to raise the inventory to the order up to level S. This policy is in general optimal when the fixed ordering costs are negligible (e.g. Scarf [17]). In most of the literature on multi–echelon inventory theory (e.g. Eppen and Schrage [6], Federgruen and Zipkin [7], Langenhoff and Zijm [14]) the order up to level policy is considered, since it was proved by Clark and Scarf [5] that in this case the inventory problem can be decomposed. Tagaras and Cohen [20] observed that low ordering costs have become increasingly common with the advent of distribution information systems, and/or the assumption that other considerations (e.g. routing schedules) dictate ordering and delivery of materials every review period. Langenhoff and Zijm [14] noted that in many practical cases the fixed costs are already accounted for at a higher level, e.g. in a multi–product environment.

If a replenishment order is placed in period t, it will be delivered in period $t+L$, i.e. there is a fixed lead time L. Let the lead time be an integral multiple of the review time, so that the review time can be taken as unity. The sequence of events in any period is: review, order, demand and delivery of the replenishment order that was placed L periods ago. Demand that cannot be satisfied from stock on hand will be backlogged, hence the ordering cost need not be considered, since all demand must be replenished and after every review an order is placed. At each review a holding cost h is charged for every unit of stock on hand, and for every unit backlogged a shortage cost p is charged. Orders arrive according to a

Poisson process with arrival rate λ, and the sizes $Y_1, Y_2, \ldots$ of the consecutive orders are i.i.d. with distribution function F. For convenience we assume that F has a continuous positive density on its support. Finally, as already mentioned in the introduction, it is assumed that the warehouse can always satisfy replenishment orders from the retailer and large orders from customers. Hence, we do not consider the inventory cost of the warehouse. We will first focus on the effects of the break quantity rule on the inventory costs at the retailer. If the break quantity rule is applied, then all orders with size smaller than or equal to the break quantity are delivered by the retailer, while the large orders are delivered from the warehouse. In the next figure the distribution system is illustrated.

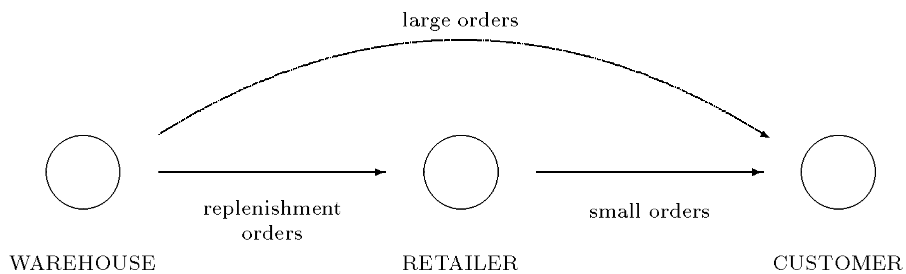

Figure 1: A two–echelon distribution system with the break quantity rule

If the break quantity equals q, the order size of the ith customer with respect to the retailer equals $Y_i 1_{\{Y_i \leq q\}}$, where

$$1_{\{Y_i \leq q\}} = \begin{cases} 1 & \text{if } Y_i \leq q \\ 0 & \text{if } Y_i > q \end{cases}$$

denotes the indicator function.

Let D_q be the total demand for the retailer during $L+1$ periods if the break quantity equals q, and let its distribution be given by G_q. Then it is well–known (see e.g. Karlin [11]) that the expected one–period holding and shortage costs for the above system are given by

$$\mathcal{L}(S, q) = h \int_0^S (S - y) dG_q(y) + p \int_S^\infty (y - S) dG_q(y) \tag{1}$$

Applying standard newsboy arguments (e.g. Porteus [16]), it follows that the optimal order up to level S^* is given by

$$S^* = G_q^{-1}\left(\frac{p}{p+h}\right)$$

with $G_q^{-1} : [0,1] \rightarrow [0,\infty)$ the inverse function of G_q, i.e.

$$G_q^{-1}(x) = \inf\{y \geq 0 : G_q(y) \geq x\}$$

If μ_q and σ_q^2 respectively denote the mean and variance of the random variable D_q, it follows that (e.g. Tijms [21])

$$\mu_q = \lambda(L+1)E[Y_i 1_{\{Y_i \leq q\}}] \tag{2}$$

$$\sigma_q^2 = \lambda(L+1)E[Y_i^2 1_{\{Y_i \leq q\}}] \tag{3}$$

Substituting S^* in (1) we obtain that the minimum expected one–period holding and shortage costs are given by

$$\begin{aligned} \mathcal{L}(S^*, q) &= -h\mu_q + (p+h)\int_{\frac{p}{p+h}}^{1} G_q^{-1}(z)dz \\ &= \sigma_q(p+h)\int_{\frac{p}{p+h}}^{1} \hat{G}_q^{-1}\, dz \end{aligned} \tag{4}$$

with $\hat{G}_q$ the distribution of the normalized random variable $\frac{D_q-\mu_q}{\sigma_q}$, i.e. $\hat{G}_q(\frac{z-\mu_q}{\sigma_q}) \equiv G_q(z)$. Unfortunately the integral in the above expression is in general difficult to compute. Moreover, it depends on the value of q. However, since $\hat{G}_q$ is the distribution of a normalized compound Poisson sum of i.i.d. random variables, we may apply the central limit theorem (Feller [8]) and approximate $\hat{G}_q$ by the standard normal distribution Φ with expectation 0 and variance 1. The main advantage of using the standard normal distribution is that for every q the distribution $\hat{G}_q$ is replaced by the same distribution, and thus the above integral is no longer depending on the break quantity.
Approximating $\int_{\frac{p}{p+h}}^{1} G_q^{-1}(z)dz$ by $\int_{\frac{p}{p+h}}^{1} \Phi^{-1}(z)dz$ it follows from (3) and (4) that

$$\mathcal{L}(S^*, q) \approx c_1\sqrt{E[Y_i^2 1_{\{Y_i \leq q\}}]} \tag{5}$$

with

$$\begin{aligned} c_1 &:= \sqrt{\lambda(L+1)}(p+h)\int_{\frac{p}{p+h}}^{1} \Phi^{-1}(z)dz \\ &= \sqrt{\lambda(L+1)}(p+h)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}(\Phi^{-1}(\frac{p}{p+h}))^2) \end{aligned} \tag{6}$$

Since $E[Y_i^2 1_{\{Y_i \leq q\}}]$ is an increasing function in q one can easily see that the expected one–period holding and shortage costs for the retailer are reduced by the break quantity rule. However, the break quantity rule also causes extra costs. Assume that the unit cost of

delivering large demand by the warehouse is given by c and the fixed costs per large order is given by K. Since the normal ordering costs are not taken into account in the analysis, the value c should be interpreted as the *additional* unit cost for not delivering the large demand from the retailer. In the value of c are included the reduction in transportation and handling costs due to the fact that the order can be delivered directly to the customer, and the penalty costs for a possible increase in waiting times for the customers with large orders. Although it is in general not easy to determine the value of c, we assume that it is known. Furthermore, observe that c does not necessarily has to be positive. In the value of K are included the possible occurrence of setup costs for delivering a large order, and a fixed penalty cost per large order. Clearly, the value of K should satisfy $K \geq 0$.
The order size of the ith customer with respect to the warehouse equals $Y_i 1_{\{Y_i > q\}}$, so the costs for serving the ith customer from the warehouse are equal to $(cY_i + K)1_{\{Y_i > q\}}$. Hence it follows by the memoryless property of the Poisson process that the expected costs for delivering the large orders separately for a typical period are given by $\lambda E[(cY_i + K)1_{\{Y_i > q\}}] = \lambda c E[Y_i 1_{\{Y_i > q\}}] + \lambda K(1 - F(q))$. These observations lead to the following conclusion.

Conclusion: *The minimum costs for the model with a break quantity* $0 \leq q \leq \infty$ *are approximated by*

$$C(q) = c_1 \sqrt{E[Y_i^2 1_{\{Y_i \leq q\}}]} + c_2 E[Y_i 1_{\{Y_i > q\}}] + c_3(1 - F(q)) \tag{7}$$

where c_1 *is given by (6),* $c_2 := \lambda c$ *and* $c_3 := \lambda K$.

By the above expression for the cost function $C(q)$ it is easy to derive a sufficient and tractable condition for the break quantity rule to be profitable, i.e. that there exists some q for which $C(q) < c_1 \sqrt{E[Y_i^2]}$. Since $E[Y_i^k 1_{\{Y_i \leq q\}}] = \int_0^q x^k dF(x)$ one can easily verify that the derivative of $C(q)$ is given by

$$C'(q) = \left(\frac{1}{2} c_1 q^2 v(q) - c_2 q - c_3\right) f(q) \tag{8}$$

where

$$v(q) := \left(E[Y_i^2 1_{\{Y_i \leq q\}}]\right)^{-1/2}$$

A sufficient condition for the break quantity rule to be profitable is given in the next theorem.

Theorem 2.1 *Assume that c_1, c_2 and c_3 are finite–valued and let $0 < u < \infty$ be given by*

$$u := \left(c_2 + \sqrt{c_2^2 + 2c_1c_3E[Y_i^2]^{-1/2}}\right)/c_1E[Y_i^2]^{-1/2} \tag{9}$$

If $F(u) < 1$, then there exists a $q < \infty$ such that $C(q) < c_1\sqrt{E[Y_i^2]}$.

Proof: Observe that $v(q)$ is decreasing in q and thus $v(q) \geq v(\infty) = \left(E[Y_i^2]\right)^{-1/2}$. Since u uniquely solves the equality

$$\frac{1}{2}c_1\left(E[Y_i^2]\right)^{-1/2}x^2 - c_2x - c_3 = 0$$

for $x > 0$, we obtain for $q > u$ that

$$\frac{1}{2}c_1v(q)q^2 - c_2q - c_3 \geq \frac{1}{2}c_1\left(E[Y_i^2]\right)^{-1/2}q^2 - c_2q - c_3 > 0$$

and so it follows by $f(q) > 0$ for all q that $C'(q) \geq 0$ for $q > u$.
Moreover, observe that $F(u) < 1$ implies that there exists a $\hat{q} > u$ such that $f(\hat{q}) > 0$, and thus $C'(\hat{q}) > 0$. Hence, there exists a $q < \infty$ such that $C(q) < C(\infty) = c_1\sqrt{E[Y_i^2]}$. □

In the next section an algorithm which determines the optimal break quantity will be presented.

3 Determining the optimal break quantity

In this section the main focus will be on solving the problem

$$\inf\{C(q) : a \leq q \leq b\} \tag{P}$$

where $C(q)$ is defined in (7), and a and b are defined as

$$\begin{aligned} a &:= \inf\{x > 0 : F(x) \geq \beta\} \\ b &:= \sup\{x > 0 : F(x) < 1\} \end{aligned}$$

with $0 < \beta < 1$. Observe that a is a lowerbound on the break quantity which is set by the management through a restriction on the percentage of customers served by the retailer. The motivation for the introduction of this restriction is threefold. First, due to fixed overhead costs at the retailer and for competitive reasons it is desirable that a large percentage of customers is served through the retailer. Secondly, it is likely that a customer ordering a small batch is less sensitive to a price reduction than a customer ordering a large batch, implying that the cost for delivering large orders from the warehouse becomes relatively higher if the break quantity decreases. Hence, the value of c and K may only hold for relatively large break quantities. Finally, the normal approximation of the lead time demand is only justified for relatively large break quantities, because for low break quantities the average number of orders in a period becomes too small and the central limit theorem may not apply anymore. Observe also that b is an upperbound on the break quantity, since by definition the size of an order cannot exceed b.

Any $q^* \in (a,b)$ solving (P) must satisfy the necessary first–order optimality condition $C'(q^*) = 0$, i.e. q^* must be a stationary point. By Theorem 2.1 we obtain that $C'(q) > 0$ for $q > u$. Since u is finite–valued it follows that the set of possible stationary points is bounded and given by $(a, \min[b,u])$. For notational convenience we define $b' := \min[b,u]$.

Before presenting an algorithm which reduces the set of possible optimal points, we need to introduce the function $\xi : [a,b'] \to [a,b']$. Let the value of $\xi(q)$ for $q \in [a,b']$ be determined by Algorithm 3.1.

Step 0 Set $i := 0$ and $q_i := q$

Step 1 Let $q_{i+1} > 0$ solve $\frac{1}{2}c_1 v(q_i) q_{i+1}^2 - c_2 q_{i+1} - c_3 = 0$,
i.e. $q_{i+1} := \left(c_2 + \sqrt{c_2^2 + 2c_1c_3v(q_i)}\right)/c_1 v(q_i)$

Step 2 If $q_{i+1} = q_i$ then $\xi(q) := q_{i+1}$ and stop,
if $q_{i+1} \le a$ then $\xi(q) := a$ and stop,
if $q_{i+1} \ge b'$ then $\xi(q) := b'$ and stop,
otherwise set $i := i+1$ and go back to **Step 1**

Algorithm 3.1: An algorithm to calculate $\xi(q)$

Observe that $\frac{1}{2}c_1 v(q) q^2 - c_2 q - c_3 = 0$ for $a < q < b'$ is equivalent to $C'(q) = 0$. Moreover, if $C'(q_i) < 0$ one can easily verify that the value of $q_{i+1} > 0$ solving $\frac{1}{2}c_1 v(q_i) q_{i+1}^2 - c_2 q_{i+1} - c_3 = 0$ must satisfy $q_{i+1} > q_i$. Since $v(q_{i+1}) < v(q_i)$ if $q_{i+1} > q_i$ it follows that $C'(q_{i+1}) < 0$, thus $q_0, q_1, \ldots$ is an increasing sequence and its limit point q_∞, if it exists, satisfies $C'(q_\infty) = 0$. A similar observation holds for the case where $C'(q_i) > 0$. We are now able to proof the next lemma that will help us to determine the set of stationary points.

Lemma 3.1 *Define* $I(x,y) := \{z : \min(x,y) \leq z \leq \max(x,y)\}$. *Then for any* $q \in [a,b']$ *it follows that* $C'(y) \neq 0$ *if* $y \in I(q,\xi(q)) \setminus \{\xi(q)\}$. *Moreover,*

$$C'(\xi(q)) = \begin{cases} \geq 0 & \text{if } \xi(q) = a \\ = 0 & \text{if } a < \xi(q) < b' \\ \leq 0 & \text{if } \xi(q) = b' \end{cases}$$

Proof: From the observations above this lemma it follows that if Algorithm 3.1 starts with $q \in [a,b']$, no point between q and $\xi(q)$ can be a stationary point, and hence the first implication is proved. If $a < \xi(q) < b$ then $\xi(q)$ is a limit point of the sequence $q_0, q_1, \ldots$ generated by Algorithm 3.1 and thus $\xi(q)$ is a stationary point. If $q_i \geq a$ and $q_{i+1} \leq a$ then either $q_{i+1} = q_i$ and thus $C'(q_{i+1}) = 0$, or $q_{i+1} < q_i$ and $C'(y) > 0$ for $q_{i+1} \leq y \leq q_i$. A similar observation holds for the upperbound b'. □

Algorithm 3.2 will reduce the set of possible optimal break quantities on $[a,b']$. It is assumed that the number of stationary points is finite. Hence, there exists an $\epsilon > 0$ such that if q_1^* and q_2^* are stationary points, it follows that $|q_1^* - q_2^*| > \epsilon$.

Step 0 Set $\mathcal{X} := \{q : a \leq q \leq b'\}$, and set $\mathcal{X}^* := \emptyset$

Step 1 Calculate $I(a,\xi(a))$ and $I(b',\xi(b'))$, set $\mathcal{X} := \mathcal{X} - I(a,\xi(a)) - I(b',\xi(b'))$ and set $\mathcal{X}^* := \mathcal{X}^* + \{\xi(a), \xi(b')\}$

Step 2 Let q be the midpoint of the largest convex subset of $\mathcal{X}$ and calculate $\xi(q)$

Step 3 Set $\mathcal{X} := \mathcal{X} - I(q,\xi(q))$ and $\mathcal{X}^* := \mathcal{X}^* + \{\xi(q)\}$

Step 4 If the length of the largest convex subset of $\mathcal{X}$ is smaller than ϵ, then go back to **Step 2**, else stop.

Algorithm 3.2: An algorithm to reduce the set of possible optimal points

Taking the element of $\mathcal{X}^*$ with the lowest associated average costs, we get the optimal break quantity. This algorithm is a special case of the C–programming algorithm introduced by Schniedovich [18]. It can easily be verified that the algorithm will terminate within $\mathcal{O}(\log \frac{b'-a}{\epsilon})$ number of iterations.

4 Computational Results

In this section, Algorithm 3.2 will be implemented to solve problem (P) for a wide variety of parameters. We will analyze the case where the order sizes of the customers are Gamma distributed. The motivation for this choice is threefold. First, for this distribution only positive order sizes are allowed, in contrast with for example the normal distribution. Secondly, the distribution can easily be fitted around the mean and the variance of the order sizes. In many practical situations managers do not have complete information on the demand distribution, but they only know the mean and variance. Finally, in order to compute $v(q)$ one only needs to compute the incomplete Gamma function, for which many numerical approximations are available. We used an approximation given by Abramowitz and Stegun [1].

In order to determine how much the approximation of the lead time demand distribution by the normal distribution influences the results, we also will calculate the exact lead time demand distribution to determine the exact optimal break quantity and its corresponding minimum costs. We used Adelson's recursion scheme (Adelson [2]) to determine the cdf of the compound Poisson distribution. However, since this recursion scheme is only defined for discrete distributions, we transformed the continuous Gamma distribution into a discrete distribution in the following way:

$$P(X = x) := \begin{cases} F(x) - F(x - 1/N) & \text{for } x = 1/N, 2/N, \ldots \\ 0 & \text{otherwise} \end{cases}$$

where $N \geq 1$ is an integer number. Clearly, the value of N determines the accuracy of the transformation. For our computations we used $N = 5$, and to determine the optimal break quantity a simple enumeration technique was used.

Generally, management will be interested in obtaining a quick estimate of the optimal break quantity. For this purpose we also analyzed the quality of approximating the optimal break quantity by u, which is given in Theorem 2.1. This value can be calculated without having to perform any numerical procedures. Henceforth we will refer to this approximation as the u–approximation.

We have analyzed the benefits of the break quantity rule for many different parameter values. However, the following parameters were always fixed: $\lambda = 10$, $L = 2$, $h = 1$, $\mu := E[Y_i] = 10$ and $\beta = 0.75$. The other parameters were varied, i.e. $p = \{5, 10, 15, 20, 25\}$, $c = \{-1, -2, 0, 1, \ldots, 7\}$, $K = \{25, 50, 75, 100\}$ and $\sigma^2 := VAR[Y_i] = \{50, 100, 150, 200, 250\}$. By varying σ^2 for μ fixed we basically vary the coefficient of variation c_Y, defined as $c_Y := \frac{\sigma}{\mu}$, which is often used as an indicator of the variability of demand. Observe that for these parameter settings the maximum coefficient of variation of the lead time demand, obtained when $\sigma^2 = 250$, is equal to $\frac{\sqrt{\sigma^2+\mu^2}}{\mu\sqrt{\lambda(L+1)}} = \frac{\sqrt{350}}{10\sqrt{30}} \approx 0.34$, and thus the probability of negative lead time demand when using the normal approximation is negligible.

In total we evaluated the effect of the break quantity rule for 1000 parameter settings. In Table 1 the cost reduction obtained by the 3 different methods is classified. In the first column the classes are presented, in the other columns the frequencies obtained by exact calculation, Algorithm 3.2, and the u–approximation are reported. We mention that the points $\xi(a)$ and $\xi(b')$, calculated in Algorithm 3.2, in all cases coincided, hence with Algorithm 3.2 the solution was quickly determined.

cost	frequencies		
reduction (%)	discrete exact calculation	Algorithm 3.2	u–approximation
[0.00, 1.00)	499	598	601
[1.00, 5.00)	165	111	109
[5.00, 10.00)	66	53	59
[10.00, 25.00)	90	75	73
[25.00, ∞)	180	163	158

Table 1: Cost reduction obtained by break quantity rule

One can see that in most of the cases the cost reduction obtained by the break quantity rule was less than 1%, but in more than 15% of the cases the reduction exceeded 25%. This implies that if the break quantity rule is profitable, the costs can be reduced considerably. We also observe that the results of the u–approximation are close to the results of Algorithm 3.2.

In table 2 some of the results are presented. We have selected the following 16 representative cases: $p = \{10\}$, $c = \{-1, 1, 3, 5\}$, $K = \{50, 100\}$ and $\sigma^2 = \{100, 200\}$. In the first four columns the values of the parameters are shown. In columns 5, 6 and 7 the optimal break quantity q^*, the corresponding percentage of customers served by the retailer $F(q^*)$, and the corresponding relative reduction in average costs $C.R.(\%)$ are presented, which are obtained by exact calculations. Note that $C.R.(\%)$ is equal to the ratio of the reduction in average costs due to the break quantity rule, and the average cost without using the break quantity rule. In columns 8, 9 and 10 similar results are presented, but then obtained by Algorithm 3.2. Algorithm 3.2 was terminated when the largest "uncovered" subset was smaller than ϵ, which we set equal to 10^{-3}. Finally, in the last three columns the value of u, the percentage of customers served by the retailer and the minimum costs are presented, when the break quantity is equal to u.

parameters				exact, $N=5$			Algorithm 3.2			u–approximation		
p	c	K	σ^2	q^*	$F(q^*)$	$C.R.(\%)$	q^*	$F(q^*)$	$C.R.(\%)$	u	$F(u)$	$C.R.(\%)$
10	-1	50	100	21	0.878	26.10	22	0.892	24.27	26	0.927	23.06
10	1	50	100	48	0.992	2.26	53	0.995	1.17	55	0.996	1.15
10	3	50	100	81	1.000	0.14	100	1.000	0.01	100	1.000	0.01
10	5	50	100	110	1.000	0.01	153	1.000	0.00	153	1.000	0.00
10	-1	100	100	37	0.975	7.70	39	0.980	6.17	41	0.984	6.06
10	1	100	100	62	0.998	0.83	69	0.999	0.32	70	0.999	0.32
10	3	100	100	90	1.000	0.07	112	1.000	0.01	112	1.000	0.01
10	5	100	100	116	1.000	0.01	161	1.000	0.00	161	1.000	0.00
10	-1	50	200	19	0.832	50.65	20	0.845	48.62	28	0.905	45.40
10	1	50	200	48	0.972	9.60	54	0.980	6.19	63	0.988	5.76
10	3	50	200	96	0.998	1.32	118	0.999	0.30	120	0.999	0.30
10	5	50	200	135	1.000	0.24	185	1.000	0.01	185	1.000	0.01
10	-1	100	200	36	0.942	24.66	38	0.948	21.70	44	0.965	20.66
10	1	100	200	66	0.990	5.13	74	0.994	2.79	79	0.995	2.72
10	3	100	200	106	0.999	0.92	131	1.000	0.18	132	1.000	0.18
10	5	100	200	142	1.000	0.19	194	1.000	0.01	194	1.000	0.01

Table 2: The effect of the break quantity rule for different methods

It can be seen that the success of the break quantity rule very much depends on the value of c. Hence, it is very important that the management has good information on the additional variable costs for delivering a large order from the warehouse. Observe that the main positive component of c is the penalty cost that has to be paid to compensate for the (possibly) longer delivery times. Therefore, the management should try to keep this cost as low as possible, e.g. by setting agreements with the customer. One can think of giving each customer with a large order some fixed price reduction if the customer accepts the longer waiting time due to the break quantity rule. For this reason, it is interesting to analyze the effect of the values of c and K on the cost reduction obtained by the break quantity rule. In Figure 1 the cost reduction is plotted for values of p and σ^2 fixed at $p = 10$ and $\sigma^2 = 150$, and for values of c and K varying from $c = [-2, 7]$ and $K = [25, 100]$. Moreover, β is set equal to zero, so no restriction on the percentage of customers served by the retailer is taken into account.

This figure is based on results obtained by using Algorithm 3.2. It shows that the sensitivity of the cost reduction for c is much larger than for K. Hence, it seems profitable to "buy off" the penalty cost for delivering large orders separately by giving a fixed price reduction for each large order. Whereas the value of c has a large influence on the fact whether or not

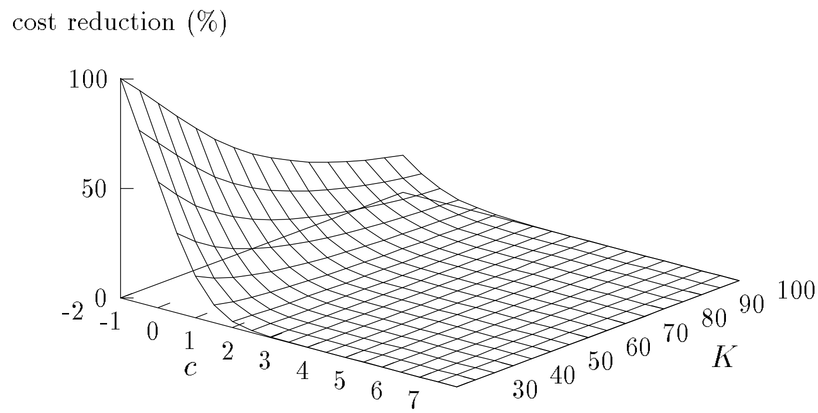

Figure 2: The impact of c and K on the cost reduction

the break quantity rule is profitable, from Table 2 it is also observed that the coefficient of variation of the order sizes has a big impact on the size of the cost reduction. For instance, for $\sigma^2 = 200$ the cost reductions are almost twice as big as for $\sigma^2 = 100$. Finally, we observed that the shortage cost p does not seem to have a significant influence on the cost reduction that can be obtained by using the break quantity rule.
The cost reduction obtained by the normal approximation of the lead time demand in all cases underestimated the actual cost reduction. Therefore, Algorithm 3.2 will generally provide a lowerbound on the cost reduction that can be obtained by using the break quantity rule.

5 Conclusions

In this paper a multi–echelon distribution system is analyzed, where orders from customers can be delivered from any level in the system. To allocate orders from customers to the different locations, we introduced a so–called break quantity rule. The break quantity determines whether an order of a customer is small or large, and the break quantity rule is implemented in the following way: a small order will be delivered from the lowest echelon, a large order will be delivered from a higher echelon. It is shown that using this rule, the inventory costs at the lowest echelon can be reduced significantly. However, in general the delivery of a large order from a higher echelon will also cause extra costs, e.g. because the

delivery lead time may increase. An approximation for the minimum costs as a function of the break quantity is derived in Section 2, and an algorithm to determine the optimal break quantity is presented in Section 3. Moreover, an easy and tractable condition under which the break quantity rule is profitable is given. This condition is that with a positive probability the order size of a customer exceeds the value u, which can be calculated without having to perform any numerical procedures.
From the computational results in Section 4 it follows that if the management can persuade the customer to accept the break quantity rule (for a possible fixed price reduction and/or small price reduction per unit demanded), and the demand is erratic, the total costs can be decreased significantly. In almost 15% of the cases we considered, the cost reduction is more than 25%. It is also shown that by using approximated minimum average cost function, the optimal break quantity and the corresponding cost reduction are well approximated. Hence, the results obtained by Algorithm 3.2 give good insight into the effect of the break quantity rule on the distribution system. Finally, it turned out that approximating the optimal break quantity by u, given in Theorem 2.1, gives satisfactory results.

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1965.

[2] R.M. Adelson. Compound Poisson distributions. *Operations Research Quarterly*, 17:73–75, 1966.

[3] S. Axsäter. Continuous review policies for multi–level inventory systems with stochastic demand. In S.C. Graves, A.H.G. Rinnooy Kan, and P. Zipkin, editors, *Handbooks in OR & MS, Vol. 4*. Elseviers Science Publishers, North–Holland, 1993.

[4] L.D. Burns, R.W. Hall, D.E. Blumenfeld, and C.F. Daganzo. Distribution strategies that minimize transportation and inventory costs. *Operations Research*, 33:469–490, 1985.

[5] A.J. Clark and H. Scarf. Optimal policies for a multi–echelon inventory problem. *Management Science*, 6:475–490, 1960.

[6] G. Eppen and L. Schrage. Centralized ordering policies in a multi–warehouse system with lead times and random demand. In L.B. Schwarz, editor, *Multilevel pro-*

duction/inventory control systems: Theory and practice. North–Holland, Amsterdam, 1981.

[7] A. Federgruen and P. Zipkin. Computational issues in an infinite–horizon, multi–echelon inventory model. *Operations Research*, 32:818–836, 1984.

[8] W. Feller. *An Introduction to Probability Theory and Its Applications, vol. 2.* Wiley, New York, 1971.

[9] G. Gallego and D. Simchi-Levi. On the effectiveness of direct shipping strategy for the one–warehouse multi–retailer R–systems. *Management Science*, 36:240–243, 1990.

[10] R. Jaikumar and V. Kasturi Rangan. Price discounting in multi–echelon distribution systems. *Engineering Costs and Production Economics*, 19:341–349, 1990.

[11] S. Karlin. Steady state solutions. In K.J. Arrow, S. Karlin, and H. Scarf, editors, *Studies in the mathematical theory of inventory and production.* Stanford University Press, Stanford, Ca., 1958.

[12] V. Kasturi Rangan and R. Jaikumar. Integrating distribution strategy and tactics: a model and an application. *Management Science*, 37:1377–1389, 1991.

[13] A.G. de Kok. Demand management in a multi–stage distribution chain. Technical Report TUE/BDK/LBS/93–35, Eindhoven University of Technology, The Netherlands, 1993.

[14] L.J.G. Langenhoff and W.H.M. Zijm. An analytical theory of multi–echelon production/distribution systems. *Statistica Neerlandica*, 44:149–174, 1990.

[15] R. Nass, R. Dekker, and W. van Sonderen-Huisman. Distribution optimization by means of break quantities: A case study. Technical Report 9375/A, Econometric Institute, Erasmus University Rotterdam, 1993.

[16] E.L. Porteus. Stochastic inventory theory. In D.P. Heyman and M.J. Sobel, editors, *Handbooks in OR & MS, vol. 2.* Elsevier Science Publishers B.V., North–Holland, 1990.

[17] H. Scarf. A survey of analytic techniques in inventory theory. In H. Scarf, D. Gilford, and M. Shelly, editors, *Multistage inventory models and techniques.* Stanford University Press, Stanford, Ca., 1963.

[18] M. Schniedovich. C–programming: an outline. *Operations Research Letters*, 4:19–21, 1985.

[19] J.E. Sussams. *Logistics Modelling.* Pitman Publishing, London, 1992.

[20] G. Tagaras and M.A. Cohen. Pooling in two–location inventory systems with non–negligible replenishment lead times. *Management Science*, 38:1067–1083, 1992.

[21] H.C. Tijms. *Stochastic models: An algorithmic approach.* Wiley, New York, 1994.