On the newsboy model with a cutoff transaction size

R. DEKKER1, J.B.G. FRENK1, M.J. KLEIJN1 and A.G. de KOK2

1 Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
E-mail: rdekker@few.eur.nl, frenk@few.eur.nl and mkl@nea.nl

2 Department of Technology Management, Eindhoven University of Technology, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
E-mail: A.G.d.kok@Tm.tue.nl

Accepted August 1998

In this paper we analyse the effect of a cutoff transaction size on the average inventory cost in a simple newsboy setting. It is assumed that customers with an order larger than a prespecified cutoff transaction size are satisfied in an alternative way, against additional cost. For compound Poisson demand with discrete order sizes, we show how to determine the average cost and an optimal cutoff transaction size. Because the computational effort to calculate the exact cost is quite large, we also consider an approximate model. By approximating the distribution of the total demand during a period by the normal distribution one can determine an expression for the average cost function that solely depends on the cutoff transaction size. A significant advantage of this approximation is that we can solve problems of any size. The quality of using the normal approximation is evaluated through a number of numerical experiments, which show that the approximate results are satisfactory.

1. Introduction

In practice, many inventory systems need to deal with erratic (or lumpy) demand patterns, which may be the result of occasionally occurring large transactions interspersed among a majority of small transactions [1,2]. An interesting concept to prevent the large transactions from disturbing the inventory system is the use of a so-called cutoff transaction size. With this concept, all demands smaller than or equal to a prespecified cutoff transaction size (or maximum issue quantity, exceptional quantity, break quantity) is routinely served, whereas large demand is alternatively delivered to the customer, e.g., by a direct delivery from a higher level stockpoint or the factory. A cutoff transaction size is easy to implement in logistics software, and it serves as a filter against very large transaction sizes (including possible entry errors). By setting the cutoff transaction size equal to infinity, or a number exceeding the maximum order size, all customer orders are delivered routinely from stock and thus the use of a cutoff transaction size can never lead to an increase in inventory cost.

If large orders are always placed by the same customers, then the management could use the concept of market segmentation and apply a separate inventory policy to each market. However, in many cases the same customer can place both small and large orders and it may be very difficult to identify different markets. Therefore, in this paper we focus on inventory systems which operate on a single market. Also in these systems the demand process may be highly variable, which is illustrated by Fig. 1. In this figure, the demand process for copper cable at a nuclear research centre in Switzerland over a period of 6 years is plotted. One can observe that the order sizes range from small (2 meter) to very large (5000 meter). At the research centre a cutoff transaction size is used to prevent the large orders from disrupting the inventory system [3]. All small orders are automatically issued from stock, whereas for each large customer order the inventory manager makes a decision. Either he/she decides to deliver the order from stock, if after satisfying the order enough inventory remains, or it is decided that the order is delivered directly from the factory. The cutoff transaction size was determined by defining a so-called exceptional threshold factor \( k \) and setting the cutoff transaction size equal to \( \mu + k\sigma \), where \( \mu \), and \( \sigma \), are respectively the mean, and standard deviation, of the order size distribution. The value of \( k \) was typically somewhere between two and five, and was manually determined by the inventory manager, without considering the cost for alternatively delivering the large orders.

There are many other examples in practice where a cutoff transaction size is used. Nass et al. [4] discuss a case study about a company who have implemented a cutoff transaction size in order to decide whether a customer order should be delivered from a local distribution centre or directly from the production plant.

The objective of this paper is to develop a method to determine the optimal cutoff transaction size, taking into account all relevant cost factors. This will be done in the
context of the well-known single period newsboy inventory model, which is the building block for a number of stochastic inventory models [5,6]. We will extend the newsboy model with the notion of a cutoff transaction size, analyse the effect on the average cost, and determine a way to obtain the optimal cutoff transaction size. We consider compound Poisson demand, and assume that the additional cost of alternatively delivering a large order (overflow cost) is known. In most practical situations, this overflow cost includes different cost factors, such as the additional transportation cost of delivering a large order directly from the factory. It may also contain negative cost factors, e.g., the reduction in transportation cost due to the fact that direct deliveries reduce the total travel distance.

In principle we are able to determine the optimal policy, but the computation times grow too large if the order sizes are significantly different. If the maximum order size exceeds 100 units, the recursive scheme for calculating the distribution of the total demand during a period explodes and the computation of the exact cost becomes impossible. A way to prevent this is to limit the problem size by making classes of order sizes and scale them down to levels between e.g., one and 100. A disadvantage of this approach is that it requires a detailed analysis of every order size distribution. A more robust approach is to approximate the distribution of the demand during a period by a normal distribution, based on the mean and variance of the distribution. We will show that the distribution of the total demand during a period by a normal distribution, based on the mean and variance of the demand is known. In most practical situations, this overflow cost includes different cost factors, such as the additional transportation cost of delivering a large order directly from the factory. It may also contain negative cost factors, e.g., the reduction in transportation cost due to the fact that direct deliveries reduce the total travel distance.

In principle we are able to determine the optimal policy, but the computation times grow too large if the order sizes are significantly different. If the maximum order size exceeds 100 units, the recursive scheme for calculating the distribution of the total demand during a period explodes and the computation of the exact cost becomes impossible. A way to prevent this is to limit the problem size by making classes of order sizes and scale them down to levels between e.g., one and 100. A disadvantage of this approach is that it requires a detailed analysis of every order size distribution. A more robust approach is to approximate the distribution of the demand during a period by a normal distribution, based on the mean and variance of the distribution. We will show that the distribution of the total demand during a period by a normal distribution, based on the mean and variance of the demand is known. In most practical situations, this overflow cost includes different cost factors, such as the additional transportation cost of delivering a large order directly from the factory. It may also contain negative cost factors, e.g., the reduction in transportation cost due to the fact that direct deliveries reduce the total travel distance.

In the next section an overview of the literature on the use of a cutoff transaction size in inventory systems is presented. In Section 3 we derive exact expressions for the average cost of a newsboy model with a cutoff transaction size. Section 4 discusses an approximate analysis, where the demand during a period is assumed to be normally distributed. In Section 5 the quality of the normal approximation is tested through a number of computational experiments. The last section provides some concluding remarks.

2. Literature review

Although the concept of a cutoff transaction size appears to gain popularity in practice it is not extensively analysed in the literature [7]. The first theoretical contribution (to the best of our knowledge) was made by Popp [8], who introduced the notion of a combined inventory policy where small demand was delivered from stock and large demand was delivered directly, thereby incurring a fixed setup cost. Using a simplified inventory model with an approximate cost function, he derived, for exponentially distributed order sizes and zero lead time, conditions under which a combined inventory policy was beneficial. Popp concluded that this was typically the case if extreme demand values exist.

Silver [18] has discussed the factors that contribute to an erratic demand pattern. First, simply there may be a large number of small customers and only a few large customers. Secondly, in a multi-echelon inventory system the erratic demand may be caused by the inventory control decisions made at the next lower echelon. It is well-known that in a multi-echelon system a non-erratic demand pattern at the lowest echelon can be turned into a highly erratic demand pattern at the higher echelons [9]. Silver [1] has introduced the notion of a cutoff transaction size and presented, for zero lead time and compound Poisson demand, a method to determine the average inventory cost and service levels in an (s,S) model with a cutoff transaction size. Basically, the method is to replace the original order size distribution by its cutoff distribution and then use the standard formulae to determine the cost and service level.

More recently, Hollier et al. [2,10] and Mak and Lai [11,12] have published a number of papers on the use of a cutoff transaction size in inventory models. Hollier et al. [2] have discussed an (s,S) inventory system with constant lead time, complete backordering and a cutoff transaction size w. All customer orders with a size exceeding the cutoff transaction size w were filtered out of the inventory system and treated as special orders to be satisfied by special deliveries, whereas all small demands were satisfied from stock. For the special deliveries of the large demands they introduced a cost function, consisting of a fixed cost for each large demand and a unit cost for each item satisfied through a special delivery. The demand distribution was approximated by a stuttering Poisson distribution (i.e., a compound Poisson distribution where the order sizes have a geometric distribution, see e.g. Adelson [13]), and they presented an algorithm to determine the optimal values of s and S for a given w. This algorithm basically replaces the original order size
distribution by its cutoff distribution, and then uses the algorithm of Zheng and Federgruen [14] to find the optimal inventory policy. The optimal value of \( w \) can be found by complete enumeration. A number of numerical examples showed that the inventory cost could be reduced by about 10%.

In a subsequent paper, Hollier et al. [10] discuss the case where this model has some slight changes. Whenever a large demand triggers a special order, this order becomes a joint replenishment order which also raises the available inventory to the order-up-to level \( S \). Furthermore, instead of a stuttering Poisson distribution, the demand is now assumed to follow a general discrete distribution by its cutoff distribution, and then uses the compound Poisson distribution. A numerical example showed that the total cost could be reduced by 7%.

Mak and Lai [12] have analysed another \((s, S)\) inventory system with a cutoff transaction size \( w \). Here, the system routinely satisfies customer orders with a size smaller than or equal to \( w \), and only supplies the cutoff amount \( w \) for customer orders with a size exceeding \( w \). Hence, the excess demand is simply refused, which causes a shortage cost and a reduction in revenues (because the selling price decreases). The demand pattern was approximated by a stuttering Poisson distribution, and an algorithm to determine the optimal \( s \) and \( S \) given the value of \( w \) (again based on Zheng and Federgruen [14]) was presented. The optimization of the cutoff transaction size \( w \) was later discussed by Mak and Lai [11]. They derived a lower bound on the average cost gain for the value of \( w \), which could be used efficiently in an enumerative search procedure for the best cutoff transaction size. A numerical example showed that the total cost could be reduced by 15%.

Finally, for an overview on the effect of using a cutoff transaction size on the overall performance of a distribution system we refer to Kleijn and Dekker [15].

### 3. Analysis of the newboy model with a cutoff transaction size

In this section, the traditional newboy model is extended with the incorporation of a cutoff transaction size. In this new model, demand from a customer is only satisfied from stock on hand if the size of the order does not exceed a prespecified cutoff transaction size, otherwise the customer is served in an alternative way, against additional cost. In order to distinguish customers by their order size, we assume that the demand is compound Poisson distributed. The notation that we use in this paper is as follows:

\[
\begin{align*}
N &= \text{(Poisson distributed) number of customers arriving during a period;} \\
\lambda &= \text{arrival rate of customers, i.e., } \lambda := \mathcal{E}(N); \\
Y_i &= \text{(random) order size of } i\text{th customer; } \\
a &= \text{order size distribution, i.e., } a(j) := \Pr(Y_i = j); \\
a(0) &= 0; \\
M &= \text{maximum order size, i.e., } M := \max\{j \geq 0 : a(j) > 0\}; \\
q &= \text{cutoff transaction size; } \\
I_A &= \text{indicator function of the event } A; \\
a_q &= \text{order size distribution for cutoff transaction size } q, \text{ i.e., } a_q(j) := \Pr(Y_i \leq q) = j; \\
D_q &= \text{(random) demand during a period for cutoff transaction size } q, \text{ i.e., } D_q = \sum_{i=1}^{N} Y_i; \\
f_q &= \text{distribution of demand during a period for cutoff transaction size } q, \text{ i.e., } f_q(j) = \Pr(D_q = j); \\
F_q &= \text{cumulative distribution of demand during a period for cutoff transaction size } q, \text{ i.e., } F_q(j) = \Pr(D_q \leq j); \\
S &= \text{order-up-to level; } \\
S(q) &= \text{optimal order-up-to level for cutoff transaction size } q; \\
C(S, q) &= \text{expected total cost during a period for order-up-to level } S \text{ and cutoff transaction size } q; \\
C(q) &= \text{minimum expected cost during a period for cutoff transaction size } q, \text{ i.e., } C(q) := C(S(q), q); \\
c &= \text{unit ordering cost; } \\
h &= \text{unit holding cost; } \\
p &= \text{unit penalty cost; } \\
\pi &= \text{overflow cost function, i.e., } \pi(j) \text{ denotes the cost of alternatively satisfying an order of size } j.
\end{align*}
\]

It is assumed that the starting inventory level is zero. Observe that the analysis can easily be adapted for situations with a positive initial inventory [5]. The conclusions, however, will remain the same and therefore we assume zero initial inventory. The problem is to determine how much to order \( S \) and how to set the cutoff transaction size \( q \), such that the expected cost over a single period is minimized. This problem reduces to the traditional newboy problem if the cutoff transaction size is set equal to infinity (or the maximum size \( M \) of a customer order), since then all demand is handled on a routine basis. The total cost consists of

- ordering or production cost for realizing the initial stock level;
- holding cost for units in stock at the end of the period;
- penalty cost for unsatisfied demand during the period;
- overflow cost for alternatively delivering large orders during the period.

It can be verified that the expected total cost for a period is given by

\[
C(S, q) = IC(S, q) + OC(q),
\]

with \( IC(S, q) \) the expected inventory (ordering, holding and penalty) cost, given by
\[ IC(S, q) = cS + h \sum_{j=0}^{S} (S-j)f_q(j) + p \sum_{j=q+1}^{\infty} (j-S)f_q(j), \]
and \( OC(q) \) the expected overflow cost which equals
\[ OC(q) = \lambda \mathbb{E}(\pi(Y_1|Y_1>q)) = \lambda \sum_{j=q+1}^{M} \pi(j)a(j). \]

Our objective is to find the solution of the optimization problem
\[ \inf\{C(S, q) : 0 \leq S < \infty, 0 \leq q \leq M\} = \inf\{OC(q) + \inf\{IC(S, q) : 0 \leq S < \infty\} : 0 \leq q \leq M\}. \]

Observe that if \( p \leq c \), the optimal order-up-to level will be zero. Hence, we will henceforth assume that \( p > c \). In order to calculate \( S(q) \) we observe that the distribution function \( f_q \) can be computed using Adelson’s recursion scheme [13]. With
\[ a_q(j) = \begin{cases} \sum_{i=0}^{j} a(i) & \text{if } j = 0, \\ a(j) & \text{if } j = 1, \ldots, q, \\ 0 & \text{otherwise.} \end{cases} \]

it follows that \( f_q(j) \) satisfies the recursive relations (see e.g., Tijms [16])
\[ f_q(j) = \begin{cases} \exp(-\lambda(1-a_q(0))) & \text{if } j = 0, \\ \lambda(j) \sum_{i=1}^{j} (j-i)a_q(j-i)f_q(i) & \text{if } j = 1, 2, \ldots. \end{cases} \]

Hence, an efficient way to determine \( S(q) \) would be to recursively calculate \( f_q(0), \ldots, f_q(j) \) until \( f_q(j) := \sum_{i=0}^{j} f_q(i) \geq (p-c)/(p+h) \). Substituting the optimal value \( S(q) \) into the cost function \( C(S, q) \) we get the one-dimensional minimum expected cost function
\[ C(q) := \inf\{C(S, q) : 0 \leq S < \infty\} = C(S(q), q). \]

Since the order sizes are discrete and bounded by \( M \), we can use enumeration over \( q = 0, \ldots, M \) to find the optimal cutoff transaction size, i.e., the optimal solution of \( \min\{C(q) : 0 \leq q \leq M\} \), and the associated expected cost. In fact, it can be verified that \( C(S, X) = C(S, q) \) with \( q = \max\{j \leq x : a(j) > 0\} \), and hence one only needs to consider cutoff transaction sizes \( q \) for which \( a(q) > 0 \). Observe that this will considerably reduce the computational effort needed to determine the optimal cutoff transaction size.

4. Approximation of the newsboy model with a cutoff transaction size

A main problem of the exact analysis is the fact that the computation time increases exponentially when the arrival rate and/or the maximum size \( M \) of a customer order increases. This problem does not occur if the total demand during a period \( D_q \) is normally distributed. Justified by the Central Limit Theorem, the normal distribution is often used as an approximation of the real demand distribution. In a recent paper, Tyworth and O’Neill [17] have reported that although this approximation in many cases leads to a miss-specification of the optimal policy parameters, the sensitivity of the expected optimal cost appears to be much less. We will now show that by approximating the distribution of the demand during a period by the normal distribution it is possible to obtain an easy expression for the minimum expected cost \( C(q) \).

First, we introduce some additional notation:
\[ \mu_q = \text{mean demand during a period for cutoff transaction size } q, \text{ i.e., } \mu_q = \mathbb{E}(D_q); \]
\[ \sigma_q^2 = \text{variance of demand during a period for cutoff transaction size } q, \text{ i.e., } \sigma_q^2 = \text{Var}(D_q); \]
\[ S_N(q) = \text{optimal order-up-to level for cutoff transaction size } q \text{ and normal demand}; \]
\[ C_N(S, q) = \text{expected total cost during a period for order-up-to level } S, \text{ cutoff transaction size } q \text{ and normal demand}; \]
\[ C_N(q) = \text{minimum expected cost during a period for cutoff transaction size } q \text{ and normal demand, i.e., } C_N(q) := C_N(S_N(q), q); \]
\[ \varphi, \Phi = \text{pdf and cdf of the standard normal distribution; } \]
\[ G_q = \text{cdf of normal distribution with mean } \mu_q \text{ and variance } \sigma_q^2. \]

For a given cutoff transaction size \( q \), it follows (see e.g., Tijms [16]) that the first two moments \( \mu_q \) and \( \sigma_q^2 \) of the variable \( D_q \) are given by
\[ \mu_q = \lambda \sum_{j=0}^{\infty} ja_q(j) = \lambda \sum_{j=1}^{q} ja(j), \]
and
\[ \sigma_q^2 = \lambda \sum_{j=0}^{\infty} j^2 a_q(j) = \lambda \sum_{j=1}^{q} j^2 a(j). \]

Approximating \( F_q \) by a normal distribution \( G_q \) with mean \( \mu_q \) and variance \( \sigma_q^2 \) it follows from (1) that the optimal order-up-to level is given by
\[ S_N(q) = \min\{j \geq 0 : j \geq \mu_q + z\sigma_q\}, \]
with \( z := \Phi^{-1}((p-c)/(p+h)) \) the safety stock multiplier. Moreover, the expected total cost during a period is approximated by
On the newsboy model with a cutoff transaction size

\[ C_N(S, q) = cS + h \int_0^S (S - y)dG_q(y) + p \int_S^\infty (y - S)dG_q(y) \]
\[ + \lambda \sum_{j=0}^M \pi(j)a(j). \]

Although the optimal order-up-to level needs to be an integer, we substitute \( S = \mu_q + z \sigma_q \) into the expected cost function \( C_N(S, q) \) to obtain that the optimal order-up-to level needs to be an integer. We substitute \( S = \mu_q + z \sigma_q \) into the expected cost function \( C_N(S, q) \) to obtain that the minimum expected cost during a period is approximately equal to (see e.g., Porteus [6]):

\[ C_N(q) = c\mu_q + \sigma_q[(c + h)z + (p + h)I(z)] + \lambda \sum_{j=0}^q \pi(j)a(j), \]
where \( I(\cdot) \) denotes the normal loss function [16,18] and \( k := (c + h)z + (p + h)I(z) \). From Tjms [16] we learn that \( I(z) = \varphi(z) - zI'(\varphi(z)) \), and since \( z = \Phi^{-1}(p - c)/(p + h) \) this implies that \( k = (p + h)\varphi(z) \). Again, we can use enumeration over all cutoff transaction sizes \( q \) for which \( a(q) > 0 \) to determine the optimal cutoff transaction size. However, for the approximate model it is possible to characterise an upper bound on the optimal cutoff transaction size.

**Lemma 1.** An optimal solution \( q^*_N \) of the optimization problem \( \inf \{ C_N(q) : 0 \leq q \leq M \} \) satisfies

\[ q^*_N \leq q_u := \max \{ j : 0 \leq j \leq M \} \]
\[ \frac{k\sigma_m^2}{n} - (\pi(j) - cj) < 0 \].

**Proof.** By (2) we obtain that

\[ C_N(q) - C_N(q - 1) = k(\sigma_q - \sigma_{q-1}) + \lambda(\pi(q) - c\sigma_q) \]
\[ > 0 \text{ for } 0 < q \leq M. \]

Due to \( (\sigma_q - \sigma_{q-1})^2 \geq 0 \) it follows that \( 2\sigma_q(\sigma_q - \sigma_{q-1}) \geq \sigma_q^2 - \sigma_{q-1}^2 \) and thus

\[ \sigma_q - \sigma_{q-1} \geq \frac{1}{2} \sigma_q^2 \text{ and thus } \]
\[ C_N(q) - C_N(q - 1) \geq \lambda a(q) \left( \frac{1}{2} k \sigma_m^2 q^2 - (\pi(q) - cj) \right). \]

Hence, since \( \sigma_q^2 \geq \sigma_m^2 \) this yields

\[ C_N(q) - C_N(q - 1) \geq \lambda a(q) \left( \frac{1}{2} k \sigma_m^2 q^2 - (\pi(q) - cj) \right). \]

Since for any \( q \geq q_u \) it follows that \( C_N(q) \geq C_N(q_u) \) the desired result follows.

An immediate consequence of the above result is that for an affine overflow cost function \( \pi(j) = \pi_0 + \pi_1j \) an upper bound on the optimal cutoff transaction order size is given by

\[ q_u = \frac{(\pi_0 - c) \sigma_M}{k} + \sqrt{\frac{(\pi_0 - c)^2 \sigma_M^2}{k^2} + \frac{2\pi_0 \sigma_M}{k}}. \]

Since this upper bound is very easy to compute, it may be used as a “quick and dirty” approximation for the optimal cutoff transaction size. This concludes our approximate analysis of the newsboy model with a cutoff transaction size.

**5. Computational results**

The main objective of this section is to test the quality of the normal approximation, in particular with respect to the optimal cutoff transaction size and the maximum cost reduction that can be obtained by introducing a cutoff transaction size. We evaluated four different order size distributions and considered affine overflow cost. For each distribution we generated examples by choosing the parameter values from the following sets:

- \( h \in \{1\} \)
- \( p \in \{10, 50, 100, 500\} \)
- \( c \in \{5, 10, 25, 50\} \)
- \( \pi_0 \in \{0, 10, 25, 100\} \)
- \( \pi_1 = c + \sigma(p - c) \) with \( \sigma \in \{0.00, 0.25, 0.50, 0.75\} \) and \( \lambda \in \{1, 2, 5, 10\} \).

Since we require \( p > c \) this leads to 768 different data sets for each distribution.

The first two order size distributions we used are based on examples given by Hollier et al. [10] and Silver [1]. The third distribution is created using a geometric distribution with parameter 0.2 \( (a(j) := 0.2 \times 0.8^{-j} \text{ for } j = 1, \ldots, 15, \text{ and setting the tail of the distribution at } j = 25) \). The last order size distribution was based on real-life demand data from the research centre discussed in Section 1 (see Fig. 1), scaled down with a factor 100 to allow for the calculation of the exact cost. The order size distributions are presented in Table 1.

In Table 2 we present the calculated cost reduction obtained by using a cutoff transaction size. The values in column Exact report the relative cost reductions obtained using the exact cost analysis, i.e.

\[ \text{Exact} := \frac{C(M) - \inf\{0 \leq q \leq M : C_N(q)\}}{C(M)} \times 100\%, \]

whereas in column Approximation the relative cost reductions are presented which were obtained using the approximate cost function, i.e.

\[ \text{Approximation} := \frac{C_N(M) - \inf\{0 \leq q \leq M : C_N(q)\}}{C_N(M)} \times 100\%. \]

We also calculated the relative cost reduction that one obtains when the cutoff transaction size equals the cutoff transaction size that minimizes the approximate cost function, i.e.

\[ \text{Optimal normal} := \frac{C(M) - \inf\{0 \leq q \leq M : C_N(q)\}}{C(M)} \times 100\%, \]

with \( q^*_N \) the optimal solution of \( \inf\{0 \leq q \leq M : C_N(q)\} \). This reduction is presented in column Optimal normal. Finally, column Upperbound reports the relative cost reduction related to using the upper bound \( q_u \) given in (3) as the cutoff transaction size, i.e.
In Table 2 the minimum, average and maximum relative cost reductions are presented for each order size distribution.

If the maximum order size or the arrival rate of customers is rather large, then it is computationally impossible to use the exact cost function $C(S, q)$ in order to determine the optimal cutoff transaction size and the corresponding relative cost reduction. In this case, one may use the approximate cost function $C_N(S, q)$. Comparing the columns Exact and Approximation, we can see how well the relative cost reduction is estimated when using the approximate cost function. From Table 2 we see that using the normal approximation of the demand generates results which are on average close to the exact results. It is also interesting to determine the quality of the optimal cutoff transaction size obtained by minimizing the approximate cost function. Comparing columns Exact and Optimal normal, we see that on average the relative cost reduction when using the approximate optimal cutoff transaction size is only slightly smaller than optimal. However, in some cases, it leads to a bad performance, i.e., an increase in cost. The worst case, for order size distribution 1, led to a cost increase of 8%. An analysis of the worst cases revealed that they occurred for $\lambda = 1$, i.e., the lowest arrival rate. Since the normal approximation is justified using the Central Limit Theorem, it is clear that its quality will increase if the arrival rate of customers increases. Using the upper bound $q_u$ as a “quick and dirty” approximation of the optimal cutoff transaction size gives satisfactory results. Although on average the relative cost reductions are less than optimal, the worst case behavior is good compared to the worst case behavior of using the cutoff transaction size that minimizes the approximate cost function. From Table 2 one can observe that the relative cost reduction appears to increase with the coefficient of variation of the order size distribution, which is defined as the ratio of the standard deviation and the mean of the order size distribution. This can be explained by the fact that the inventory holding cost is increasing with the variability of the demand. If this variability is relatively large, then rejecting the demand from a small fraction of customers with large order sizes will cause a significant reduction in the demand variability and thus the inventory holding and shortage cost.

So far, we have only presented average results about the quality of the normal approximation. However, for a particular data set the difference between the approximated and exact cost reduction may be much larger than the average difference. Therefore, we have calculated for each data set the difference between the exact cost reduction and the other cost reductions, i.e.,

<table>
<thead>
<tr>
<th>Table 1. Order size distributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order size distribution 1 (mean 4.49, variance 35.48, coef. of var. 1.32)</td>
</tr>
<tr>
<td>$i$</td>
</tr>
<tr>
<td>$a(i)$</td>
</tr>
<tr>
<td>$i$</td>
</tr>
<tr>
<td>$a(i)$</td>
</tr>
<tr>
<td>$i$</td>
</tr>
<tr>
<td>$a(i)$</td>
</tr>
</tbody>
</table>

| Order size distribution 2 (mean 2.1, variance 54.99, coef. of var. 3.53) |
| $i$ | 1 | 5 | 75 |
| $a(i)$ | 0.90 | 0.90 | 0.01 |

| Order size distribution 3 (mean 5.16, variance 25.01, coef. of var. 0.96) |
| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $a(i)$ | 0.20 | 0.16 | 0.128 | 0.102 | 0.082 | 0.066 | 0.052 | 0.042 | 0.034 | 0.027 |
| $i$ | 1 | 11 | 12 | 13 | 14 | 15 | 25 |
| $a(i)$ | 0.022 | 0.017 | 0.014 | 0.011 | 0.009 | 0.003 |

| Order size distribution 4 (mean 11.16, variance 127.73, coef. of var. 1.01) |
| $i$ | 1 | 2 | 3 | 5 | 6 | 7 | 8 | 10 | 11 | 12 |
| $a(i)$ | 0.18 | 0.02 | 0.08 | 0.04 | 0.14 | 0.02 | 0.02 | 0.10 | 0.14 | 0.02 |
| $i$ | 13 | 16 | 18 | 20 | 21 | 22 | 30 | 35 | 38 | 46 |
| $a(i)$ | 0.02 | 0.02 | 0.02 | 0.04 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 |
| $i$ | 50 |
| $a(i)$ | 0.02 |
determined $\text{Approximation-dif} := \text{Exact} - \text{Approximation}$, $\text{Optimal Normal-dif} := \text{Exact} - \text{Optimal normal}$, and $\text{Upperbound-dif} := \text{Exact} - \text{Upperbound}$. The minimum, average and maximum values are reported in Table 3. Observe that the difference between the exact cost reduction and the reduction obtained by using the normal approximation can be very large. For order size distribution 2, the normal approximation may overestimate the cost reduction by 55%. On the other hand, for the same distribution the exact cost reduction may be underestimated by 62%. Both extreme cases occurred for $\lambda = 1$, $p = 500$ and $x = 0.75$. For this parameter combination the normal approximation performs worst, and the cost error is high due to the high values of the cost parameters.

Intuitively, one could imagine that there is a relation between the optimal cutoff transaction size and the order size distribution. In particular, it is expected that an optimal cutoff transaction size will coincide with a peak in the distribution. However, this does not seem to be the case in general. For all order size distributions we determined the percentage of cases that a certain cutoff transaction size was optimal, and compared these with the order size distributions. We have not found a relation between the “popular” optimal cutoff transaction sizes and peaks in the order size distributions.

We now consider an arbitrary example for order size distribution 4. We have plotted in Fig. 2 the exact cost function $C(q)$ and the approximate cost function $CN(q)$ for all values of the cutoff transaction size $q$.

Even though there is a significant difference between the exact and approximated cost in Fig. 2, the shapes of the exact and approximate cost function are similar and the optimal cutoff transaction sizes are close (18 (Exact) versus 13 (Approximation)), and also the relative cost reductions are similar (7 versus 10%). The value of the upper bound $qu$ in Fig. 2 was equal to 27, which led to a reduction of 6% in average cost.

To conclude our section on computation results, we observe from Fig. 2 that in general the cost functions $C(q)$ and $CN(q)$ do not have a shape that could allow us to design a straight algorithm to find the optimal $q$. Together with the observation that the optimal cutoff transaction size does not seem to have a clear relation with the order size distribution, this justifies the use of enumeration to find the optimal cutoff transaction size. The remark at the end of Section 3, that only cutoff transaction sizes $q$ for which $a(q) > 0$ need to be evaluated, can also be verified from Fig. 2. One may observe that $C(q + 1) = C(q)$ and $CN(q + 1) = CN(q)$ if $a(q + 1) = 0$.

6. Concluding remarks

In this paper an analysis of the newsboy model, extended with the inclusion of a cutoff transaction size, was pre-
This extension allows for the delivery of large demands in an alternative way, thus preventing the large orders from disrupting the inventory system. The main contributions of this work are the derivation of the exact cost and an approximate expression of the cost solely as a function of the \( c_{\text{opt}} \) transaction size. The approximation originates from fitting a normal distribution on the distribution of the total demand during a period. From the computational experiments it follows that this approximate analysis gives satisfactory results. A major advantage of using the normal approximation is the fact that it requires much less computational effort. Therefore, it can handle order size distributions with a wide range of possible order sizes, whereas the computational effort needed to calculate the exact cost grows exponentially with the range of the order size distribution. The results presented in Section 5 indicate that the optimization problem associated with finding the optimal cutoff transaction size is in general not an easy problem due to the non-convexity of the average cost function. Since only a relatively small number of cutoff transaction sizes need to be evaluated, the use of enumeration to find the optimal policy is justified. It is possible to extend the approximate results to more general systems, since the newsvendor type equations appear in many inventory models. For example, the results can be extended to a multi-period, multi-echelon inventory system with positive lead times [19].

Finally, we mention that whilst we have analysed the case of a newsvendor model with a stationary cutoff transaction size, it is also possible to implement more sophisticated rules. For example, the cutoff transaction size could be dependent on the inventory level, or on the remaining lead time until the next incoming order. As such, the analysis in this paper is closely related to inventory rationing problems (see e.g., Topkis [20]). However, a stationary cutoff transaction size is still useful in practice, since it can be used to identify a large order. The inventory manager should then decide whether or not to deliver the large order from stock on hand.

### Acknowledgement

The authors thank the anonymous referees for their valuable comments which helped to improve this paper.

### References


### Table 3. Minimum, average and maximum difference between exact and approximate cost reductions

<table>
<thead>
<tr>
<th>Order size distribution</th>
<th>Approximation-dif</th>
<th>Optimal normal-dif</th>
<th>Upperbound-dif</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum (%)</td>
<td>Average (%)</td>
<td>Maximum (%)</td>
</tr>
<tr>
<td></td>
<td>Minimum (%)</td>
<td>Average (%)</td>
<td>Maximum (%)</td>
</tr>
<tr>
<td></td>
<td>Minimum (%)</td>
<td>Average (%)</td>
<td>Maximum (%)</td>
</tr>
<tr>
<td>1</td>
<td>-23</td>
<td>-0.1</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>-55</td>
<td>-5.4</td>
<td>62</td>
</tr>
<tr>
<td>3</td>
<td>-16</td>
<td>0.1</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>-16</td>
<td>0.0</td>
<td>18</td>
</tr>
</tbody>
</table>

Fig. 2. A plot of \( C(q) \) and \( C_N(q) \), for distribution 4 and \( c = 5 \), \( p = 10 \), \( n_0 = 25 \), \( n_1 = 6 \) and \( \lambda = 5 \).
On the newsboy model with a cutoff transaction size


Biographies

Rommert Dekker obtained his M.Sc. and Ph.D. degrees in Mathematics from the University of Leiden, The Netherlands (’81 and ’85) and his M.Sc. degree in Industrial Engineering from the Twente University of Technology (’89). From 1985 till 1992 he worked with Shell Research and Shell International in the areas of maintenance optimization and refinery logistics. Since 1992 he has been a full professor of Operations Research at the Faculty of Economics, Erasmus University Rotterdam. He has done several case studies in logistics and maintenance and published over 50 papers in international journals. His present research interests include inventory control, distribution optimization and reverse logistics.

Hans Frenk obtained his M.Sc. and Ph.D. degree in Mathematics from Utrecht University (’78) and the Erasmus University Rotterdam (’82). Presently he is associate professor at the latter university. His research interests include nonlinear programming, convex analysis and stochastic models.

Marcel Kleijn received his M.Sc. and Ph.D. in Operations Research from the Erasmus University Rotterdam (’94 and ’98). His thesis concerned demand differentiation in inventory control. Presently he is working as a transport economics consultant for NEA in the Netherlands.

Ton de Kok graduated at the University of Leiden in Mathematics and Economics in 1981 and attained the doctorate in 1985 from the Free University of Amsterdam. Thereafter he joined the Centre for Quantitative Methods (CQM) of Philips Electronics as consultant in Operations Research. During 5 years he participated in over 20 client projects in the field of logistics and production management. In 1990 he was appointed Manager Logistics Innovation of Philips Consumer Electronics. Since 1992 he has been a full professor of Operations Management at the faculty of Technology and Management at the Eindhoven University of Technology. His main research areas are Production Management and Supply Chain Management with emphasis on quantitative analysis. In this area he published over 40 papers in international and national journals.