The Newsboy Problem with Resalable Returns

Julien Mostard* Ruud Teunter†

August 7, 2002

Abstract

We analyze a newsboy problem with resalable returns. A single order is placed before the selling season starts. Purchased products may be returned by the customer for a full refund within a certain time interval. Returned products are resalable, provided they arrive back before the end of the season and are undamaged. Products remaining at the end of the season are salvaged. All demands not met directly are lost. We derive a simple closed-form equation that determines the optimal order quantity given the demand distribution, the probability that a sold product is returned, and all relevant revenues and costs. We illustrate its use with real data from a large catalogue/internet mail order retailer.

Keywords: inventory, newsboy problem, product returns, reverse logistics, mail order retailer

1 Introduction

In many businesses, customers have the legal right to return a purchased product within a certain time frame. The money is then partially or fully reimbursed and the product can be resold if the quality is good enough and there still exists demand for it.

*Erasmus University Rotterdam, Rotterdam School of Management, PO Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: jmostard@fbk.eur.nl

†Erasmus University Rotterdam, Econometric Institute, PO Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: teunter@few.eur.nl

For several reasons, such returns are especially apparent in catalogue/internet mail order companies. First of all, customers buy via a catalogue or a portal such as the internet and thus do not get to see the physical product before making their purchase decision. Consequently, the product often turns out to be the wrong size or shape or the color differs slightly from that shown in the catalogue or on the internet. Second, the main attractiveness of this ‘distant shopping’, being the ease with which one can order products from home without going anywhere, simultaneously constitutes its main downside. It is easy to return products. After filling out a return form, the product is collected or can be returned via mail. The purchase price and often also the shipping costs are refunded. Moreover, the fact that you do not have to bring a product back personally makes the return process anonymous. For the catalogue/internet mail order retailer that motivated this study (see also the next section), return rates can be as high as 75%.

Since returned products can be resold (if they are undamaged and returned before the end of the selling season), returns should be taken into account when taking ordering decisions. In this paper, we will show how this can be done for the case that a single order is placed for each product. Of course, this order should arrive before the start of the season. We note that this single order case is not unrealistic. Large ordering lead times (e.g. due to production in south-east Asia) and short selling seasons (one summer or one winter) often force mail order retailers to order the entire collection long before the start of the season.

So, we analyze the problem of determining the optimal order quantity for a single order, single period problem. This problem is well-known as the *newsboy problem* or the *news vendor problem*, and has been studied extensively in the literature. See Silver et al. Khouja [1] and [2] for overviews. However, to the best of our knowledge, Vlachos and Dekker [3] are the only ones who include a return option.

The analysis of Vlachos and Dekker [3] is based on two very restrictive assumptions. The first is that a fixed percentage of sold products will be returned. As a result, part of the variability in the net demand is ignored. Since demand variability is a key factor in the analysis, this leads to sub-optimal ordering quantities. The second restrictive assumption is that products can be resold at most once. But if return rates are high, it is likely that products are resold more than once. Indeed, that often occurs at the mail order retailer that will be discussed in the next section.

In this paper we analyze the newsboy problem with resalable returns, but without these restrictive assumptions. Each sold product is returned with a certain probability. Products can be resold any number of times. We derive a simple formula that determines the optimal ordering quantity. Using real data of the mail order retailer, we illustrate the use of this formula for a large selection of products from a certain selling season. Furthermore, we compare the resulting order quantities to those that were proposed by Vlachos and Dekker [3] and to the orders the company would have placed using its ordering rule.

The remainder of this paper is organized as follows. In Section 2, we describe the case study that motivated this research. In Section 3, we present the mathematical model and discuss the assumptions. Section 4 reviews the approximate analysis of Vlachos and Dekker [3]. The exact analysis is presented in Section 5. Section 6 discusses the available data from the case study, which we use to illustrate the results. Our procedure for estimating the mean and variance of gross demand for every product is shown in Section 6.1. For our computations, we assume Normality of gross demand. In Section 6.2, we show that, given this assumption, the distribution of net demand is approximately Normal for all products in our data sets. Section 7 describes the computational experiments and in Section 7.1 we show and analyze the cumulated results over all products as well as the detailed results for a selected group of nine products. Finally, we summarize our findings and indicate directions for further research in Section 8.

2 Case study

This research is motivated by a case study at a large mail order retailer. This company sells a broad range of hardware and fashion products via a catalogue and to a lesser extent via the internet. By law, customers have the right to return products within 10 days after delivery. In practice, however, the company allows returns after this period and the bulk of returns arrives in the second and third week after delivery.

This research concentrates on the fashion products, since these all have a single selling season and involve high return rates. The return rates are generally around 35% to 40%

but can be as high as 75%. Products can be returned for free and are collected at the customers' homes, which is the main explanation for the high return rates.

There are two selling seasons, summer and winter, which both last 26 weeks. The manufacturers are situated in south-east Asia, and the order lead time is large (up to 14 weeks, including product development time). Two 'regular' orders are placed before the start of the season, and sometimes a third 'emergency' order is placed during the season.

Due to the large lead time, the first regular order is placed long before the season starts, so that it arrives in time. At the time of placing the first regular order, only a rough prognosis of season demand is available. Therefore, the first order is small (not more than 60% of the prognosis) so that the risk of ordering too much is minimal.

Between placing the first and the second regular order, demand information is gathered by sending a selection of loyal customers a preview catalogue and allowing them to place orders immediately. That leads to a better estimate, the so-called *preview*, of season demand. The second regular order, based on the preview, arrives in the start of the season, approximately three weeks after the start.

If demand during the first weeks of the season is much higher than expected, then an additional emergency order is sometimes placed. The lead time associated with that order is minimized by transporting from the regular (south-east Asian) manufacturer via air instead of road/sea, or by using a regional (eastern European) manufacturer. Of course, emergency ordering does lead to higher purchase and ordering costs.

It is clear from this description, that many of the characteristics of this case study fit the general situation outlined in Section 1. There is one short selling season, order lead times are large, and return rates are high. But an important difference is that, instead of a single order, two and sometimes even three orders are placed. Hence, the problem of determining optimal orders is more complex than the newsboy problem.

However, the key problem is that of determining the optimal second regular order (up-to level). The first regular order is small, and only needed to cover the period until the second order arrives. Emergency ordering is expensive, and should therefore be avoided as much as possible. By focussing only on the second regular order, the ordering problem reduces to the newsboy problem. So we can apply the newsboy framework of this paper to get insights into the optimal second regular order. We will do so in Sections 7 and 7.1

using real data.

3 Model and assumptions

There is a single replenishment opportunity at which Q products are ordered. Those products arrive before the start of the selling season. The total number of products demanded during the season, i.e. gross demand, is denoted by G. Its mean and standard deviation are denoted by μ_G and σ_G, respectively. Customers are allowed to return purchased products within a certain time limit (usually 7-30 days in practice). If a product is returned, the customer gets a full refund. We assume that each sold product is returned with the same probability r.

An undamaged returned product is *resalable*, but only if it is returned/collected, inspected/tested, and put back on the shelf before the selling season ends. Moreover, there must be sufficient demand to sell the returned product (*assuming priority of resales over first sales*). We remark that this priority assumption is only needed to classify a returned product as resalable or not, i.e. for defining a resalable return. The priority rule does not effect the profit, since returned products are as-good-as-new and sold at the same price as new products. So, the model is valid for any priority rule used.

It is clear from the above, that in order to determine the probability that a product is resalable if it is returned, we need to know the selling date, the distribution of the time between a sale and a return, the collection and test times, and the demand curve (information on the total season demand is insufficient). To avoid the need for all this information and keep the analysis tractable (and practical), however, we assume that the probability that a product is resalable if it is returned is fixed and known, and denote it by k.

We remark that for the practical case of the mail order retailer, the average time between a sale and a return plus the collection and test times is about 2-3 weeks and hence small relative to the length of the selling season (26 weeks). The (expected) number of demands is larger than the number of returns during almost all of the season (except the last 4 weeks) for all products. So, almost all returns that are back on the shelf before the season ends are indeed resalable. These characteristics justify the assumption of a fixed (high) probability that a return is resalable.

The objective is to find the order quantity Q that maximizes the expected profit. The relevant revenue and cost parameters (all per unit) are the selling price p, the salvage value s, the purchase cost c, the (gross demand) shortage/loss of goodwill cost g, and the collection cost d.

We introduce some more notations that will appear to be useful in the analysis that follows. Let 'net demand' N denote the total number of (gross) demanded products that are not returned *and* resalable (either not returned or returned but not resalable), assuming that all demands are met. Its mean and standard deviation are denoted by μ_N and σ_N, respectively. Let $p_G = (1-r)p - rd + r(1-k)s$, which can be interpreted as the unit expected revenue of satisfying a gross demand, including salvage revenue if the sold product is returned but not resalable. Let $p_N = (1 + rk + (rk)^2 + \ldots)p_G = p_G/(1-rk)$, which can be interpreted as the unit expected revenue of satisfying a net demand, i.e. of (repeatedly) selling a product until it is not returned and resalable, including salvage revenue if the sold product is returned but not resalable. Let $g_N = g/(1-rk)$ denote the expected net shortage cost of not satisfying a net demand.

The notations that have been introduced and some additional notations that will be used in the remainder are listed in Table 1.

INSERT TABLE 1 ABOUT HERE

4 An approximation of the optimal order quantity

Before we determine the exact optimal order quantity in the next section, we first present and discuss an approximation that was proposed by Dekker and Vlachos [3]. They studied the same model as we do. However, their analysis is based on two additional assumptions that will be given below. So the resulting ordering quantity is only approximately optimal. In a later section, we will study its performance for a number of real life examples.

The two simplifying assumptions are as follows.

- *Products can only be resold once.* The authors defend this assumption by remarking that products are generally resold at the end of the selling season and do not return again before the end of the season.

- *A fixed percentage r of sold products is returned (and resalable).* So if n products are sold, then exactly rn of those products are returned of which exactly krn products are resalable. With this assumption, part of the variability in the number of (resalable) returns, given gross demand, is ignored.

It is not clear, in advance, what the joint effect of these two assumptions on the order quantity is. Ignoring part of the variability in the net demand will lead to a smaller 'safety stock' and hence to an underestimation of the optimal order quantity. But assuming that products can only be resold once will lead to an overestimation of the optimal order quantity, as is illustrated by the following simple *deterministic* example. Assume that 300 products are (gross) demanded, every second sold product is returned ($r = 0.5$), and all returns are resalable ($k = 1$). Then, given our assumption that the demand rate is always larger than the resalable return rate (see the previous section), an order for 150 products is sufficient to meet all demands. But under the assumption that products can only be resold once, an order for 200 products is needed.

Under the above two assumptions, it is easy to determine the optimal order quantity. We will illustrate this for the case with continuous demand, but the approach can also be used for cases with discrete demand. Recall that for continuous demand cases, the density function and distribution function of gross demand are denoted by f_G and F_G, respectively.

Note that the two above assumptions do not change the expected gross revenue $p_G = (1-r)p-rd+r(1-k)s$, which includes the collection cost if a product is returned and the salvage revenue if a product is returned but not resalable. The remaining salvage revenues are those for products that are never sold and for products that are returned and resalable but not resold (if priority is always given to resales over first sales, then only products that are never sold remain). So, the assumptions lead to the following approximation of the total expected profit

$$\begin{aligned}\widehat{EP}(Q) &= p_G(\mu_G - \widehat{ES}_G(Q)) - cQ - g\widehat{ES}_G(Q) + s(Q - (1-rk)(\mu_G - \widehat{ES}_G(Q))) \\ &= (p_G - s(1-rk))\mu_G - (c-s)Q - (p_G - s(1-rk) + g)\widehat{ES}_G(Q), \qquad (1)\end{aligned}$$

where $\widehat{ES}_G$ denotes the expected gross shortage, i.e., the expected number of gross demands not met. Under the two above mentioned additional assumptions, it is easy to see that at most $Q(1+rk)$ (gross) demands can be met. So the expected shortage is approximated by

$$\begin{aligned}\widehat{ES}_G(Q) &= E[G - Q(1+rk)]^+ \\ &= \int_{Q(1+rk)}^{\infty} (x - Q(1+rk)) f_G(x)dx \\ &= \int_{Q(1+rk)}^{\infty} x f_G(x)dx - Q(1+rk)(1 - F_G[Q(1+rk)]),\end{aligned}$$

which gives

$$\frac{d\widehat{ES}_G(Q)}{dQ} = -(1+rk)(1 - F_G[Q(1+rk)]). \qquad (2)$$

Combining (1) and (2) gives

$$\frac{d\widehat{EP}(Q)}{dQ} = -(c-s) + (p_G - s(1-rk) + g)(1+rk)(1 - F_G[Q(1+rk)]).$$

The order quantity resulting from this approach is therefore

$$\widehat{Q} = \frac{1}{1+rk}\, F_G^{-1}\left(\frac{(p_G - s(1-rk) + g)(1+rk) - (c-s)}{(p_G - s(1-rk) + g)(1+rk)}\right). \qquad (3)$$

Vlachos and Dekker [3] derive the same result using a slightly different analysis.

5 The exact optimal ordering quantity

Recall that the approximately optimal ordering quantity, derived in the previous section, was based on two simplifying assumptions. Those assumptions were needed, because the focus was on gross demand rather than net demand. In this section, we will not use any simplifying assumptions, and use a 'net demand approach' to determine the exact optimal ordering quantity.

Recall that the expected net revenue p_N includes the collection cost if a product is returned and the salvage revenue if a product is returned but not resalable. The remaining salvage revenues are those for products that are never sold and for products that are returned and resalable but not resold (if priority is always given to resales over first sales, then only products that are never sold remain). Hence, in 'net terms', the total expected profit can be expressed as

$$
\begin{aligned}
EP(Q) &= p_N(\mu_N - ES_N(Q)) - cQ - g_N ES_N(Q) + s(Q - (\mu_N - ES_N(Q))) \\
&= (p_N - s)\mu_N - (c - s)Q - (p_N - s + g_N) ES_N(Q) \ , \qquad (4)
\end{aligned}
$$

where ES_N denotes the expected net shortage, i.e., the expected number of net demands not met.

Since the maximum number of net demands that can be fulfilled is Q, the expected net shortage is

$$
\begin{aligned}
ES_N(Q) &= E[N - Q]^+ \\
&= \sum_{l=Q+1}^{\infty} (l - Q) \Pr[N = l]
\end{aligned}
$$

and hence

$$
\begin{aligned}
ES_N(Q) - ES_N(Q-1) &= -\Pr[N \geq Q] \\
&= -(1 - \Pr[N < Q]) \qquad (5)
\end{aligned}
$$

Combining (4) and (5) gives

$$
EP(Q) - EP(Q-1) = -(c - s) + (p_N - s + g_N)(1 - \Pr[N < Q])
$$

So the optimal order quantity Q^* is the largest value of Q for which

$$
\Pr[N < Q] > \frac{(p_N - s + g_N) - (c - s)}{(p_N - s + g_N)},
$$

i.e. the largest value of Q for which

$$
\sum_{n=0}^{\infty} \sum_{m=n-Q+1}^{n} \Pr[G = n] \binom{n}{m} (rk)^m (1 - rk)^{n-m} > \frac{(p_N - s + g_N) - (c - s)}{(p_N - s + g_N)} \ .
$$

In practice it will be difficult, if not impossible, to estimate all probabilities $\Pr[G = n]$, $n = 0, 1, 2, \ldots$. An easier alternative is to estimate the mean and the variance of gross demand, and then fit a continuous (e.g. Normal) distribution. Using the below expressions for the mean and the variance of net demand, which are proven in appendix A, the same can be done for net demand.

$$\mu_N = (1 - rk)\mu_G \quad \text{and} \quad (\sigma_N)^2 = (1 - rk)^2(\sigma_G)^2 + rk(1 - rk)\mu_G \tag{6}$$

Note that $(\sigma_N)^2 > (1 - rk)^2(\sigma_G)^2$. This is in correspondence with our statement in Section 4 that assuming fixed return and resalable rates leads to an underestimation of the variability in the number of resalable returns, and hence to an underestimation of the variability in the net demand.

Denoting the continuous distribution function of net demand by $F_N(.)$, we then get

$$Q^* = F_N^{-1}\left(\frac{(p_N - s + g_N) - (c - s)}{(p_N - s + g_N)}\right). \tag{7}$$

In our computational experiments using real data in Section 7, we will assume Normality of both gross demand and net demand. Assuming Normality of gross demand is common practise. As will be shown in Section 6.2, the distribution of net demand is approximately Normal, if the distribution of gross demand is Normal.

6 Data

For our computations, we use data provided by a large mail order company. See Section 2 for a short description of the company. Recall from that section that we focus (for each product) on the order that is placed after the previews of season demand and of the return probability become available, and consider this to be the only (newsboy) order that is placed. So, we will restrict the discussion to data that is relevant for placing this order.

There are two data sets. Set 1 consists of 4761 products, for which the preview of gross demand and the realized gross demand are given. Set 2 consists of 427 products, and *additionally* gives the preview of the return rate/probability, the realized return rate, the salvage value, the purchase cost, the return collection cost, and the sales price. Table 2 gives a quick impression of the ranges of the parameters.

INSERT TABLE 2 ABOUT HERE

Most information is available for the products in Set 2. In Section 7, we shall determine the order quantities $\widehat{Q}$ and Q^* for those products. However, some relevant information for doing so is missing: the probability k that a return is resalable, the shortage (loss of goodwill) cost g, and the distributions of gross and net demand. After discussions with the retailer, we decided to set $k = 0.95$ for each product. These discussions also convinced us to use the same shortage cost for each product (since every shortage results in a dissatisfied customer), though no indication was given about the right value. In our computations, we will try and compare different values of the shortage cost. In the remainder of this section, we describe how demand distributions are estimated.

First, in Section 6.1, we derive estimators for the mean and variance of gross demand using the data in Set 1. Note that via (6), these estimators can also be used to obtain estimates for the mean and variance of net demand. Then, in Section 6.2, we argue that it is reasonable to assume that both gross and net demand are Normally distributed.

6.1 Estimating the mean and variance of gross demand

In this section, we propose estimators for the mean and the variance of gross demand for a product, given only the preview estimate of mean gross demand. These estimators are based on the data (demand preview and demand realization) in Set 1. Since this data concerns a single selling season, it is not possible to do a time series analysis for each product separately. Instead, we obtain estimators for the mean and the variance of gross demand by combining the data for all products. To avoid additional notation, those estimators are simply denoted by μ_G and $(\sigma_G)^2$, respectively. The preview of mean gross demand is denoted by μ_P.

We restrict our attention to estimators with the following simple structure:

$$\mu_G = a\mu_P, \tag{8}$$

$$(\sigma_G)^2 = b(\mu_G)^c, \tag{9}$$

where a, b, and c are constants. These constants will be based on the data in Set 1.

Recall that Set 1 contains 4761 products, for which the preview and the realization of gross demand are given. However, we remove all 824 products with a preview of less than 150. For those products, the preview is very unreliable (often more than a factor 10 wrong, e.g. preview 4 and realization 112). So, the reduced data-set contains 3937 products with a preview of at least 150. The average of a certain expression E over these 3937 products from Set 1 is denoted by $\text{AVERAGE}_1(\text{E})$.

In order to get an unbiased estimator μ_G, we compute the ratio of realized gross demand over the preview, G/μ_P, for all products. This leads to the following parameter value for a.

$$a = \text{AVERAGE}_1\left(\frac{G}{\mu_P}\right) = 0.856.$$

Note that this implies that, on average, the realization of gross demand, G, is 14.4% lower than the preview. Even after discussions with the retailer, the reason for this considerable bias remains unclear.

We continue with the variance estimator $(\sigma_G)^2$. This estimator is determined by the two parameters b and c. For a fixed value of c, it is logical to set

$$b = \text{AVERAGE}_1\left(\frac{(G-\mu_G)^2}{(\mu_G)^c}\right). \tag{10}$$

So we will restrict our attention to c, and calculate the corresponding value of b using the above equation. The value of c (and the corresponding value of b) should be such that (9) approximately holds for the entire range of values for μ_G in our data. We therefore divide the 3937 products into 10 groups according to increasing ranges of μ_G, and look for that value of c for which $(G-\mu_G)^2/(\mu_G)^c$ is 'most constant'. It appears from Table 3 that $(G-\mu_G)^2/(\mu_G)^c$ increases with the average value of μ_G for $c \leq 1.5$, decreases for $c \geq 1.9$, and is reasonably constant in between. We therefore set $c = 1.7$. The corresponding value for b, determined by (10), is 1.84.

INSERT TABLE 3 ABOUT HERE

Since Table 3 only shows summarized data, we also present a scatter-plot of $(G - \mu_G)^2/(\mu_G)^{1.7}$ against μ_G for all 3937 products in Figure 1. It is important to remark that the seemingly decreasing pattern in this figure is misleading. It is caused by the fact that the bulk of products has a relatively low μ_G, and hence most high-valued extremes occur for low values of μ_G.

INSERT FIGURE 1 ABOUT HERE

To summarize, we propose the following estimators for the mean and the variance of gross demand:

$$\mu_G = 0.856\mu_P, \tag{11}$$

$$(\sigma_G)^2 = 1.84(\mu_G)^{1.7}. \tag{12}$$

In Section 7, we will apply these for the products in Set 2.

6.2 Normality of net demand

Since the data on gross demand is limited, it is impossible to compare different distributions with respect to their 'fit'. In our numerical experiments in the next section, we will therefore assume that gross demand is Normal. In this section, we show that under that assumption, net demand is also approximately Normally distributed. We do so for three products (numbered I, II, and III in this section) from Set 2, but the other products in this set produce similar results.

Table 4 gives the relevant characteristics of the three considered products. The estimated mean and standard deviation of gross and net demand follow from (6), (11), (12), and setting $r = r_P$ (the return rate preview is not significantly biased) and $k = 0.95$ (see Section 6). Products I, II, and III in Table 4 are selected from Set 2 such that the ratio μ_N/σ_N is small, medium, and large, respectively.

INSERT TABLE 4 ABOUT HERE

The procedure for finding the distribution of net demand, under the assumption that gross demand distribution is Normal with mean μ_G and standard deviation σ_G, is as follows. We draw from the Normal gross demand distribution until per product 1000 positive random drawings $g_i, i = 1, 2, \ldots, 1000$ are obtained (negative drawings are left out). For each g_i, the corresponding number of resalable returns r_i is drawn (once) from a binomial distribution with g_i repetitions and probability of success $0.95r_P$ (recall that k is set to 0.95 for all products). Computing $n_i = g_i - r_i$ then gives 1000 random values of net demand per product. The net demand distribution is obtained by assigning probability 1/1000 to each n_i.

Figures 2a-2c compare the distribution function of net demand to that of a Normal distribution with mean μ_N and standard deviation σ_N (determined by (6)) for products 1-3. It appears that for all three products, the net demand distribution is close to Normal. In fact, the difference between the distribution functions is not caused by the non-Normality of net demand if gross demand is Normal. Instead, it results from the non-Normality of gross demand, since demands cannot be negative. This explains why the difference decreases with μ_N/σ_N. Fortunately, that ratio is higher than 1.5 for all products that will be considered in the next section, justifying the assumption of Normal gross demand.

INSERT FIGURES 2a-2c ABOUT HERE

7 Computational Experiments

Using the data in Set 2, we perform computational experiments to compare the exact optimal order quantity Q^* to the approximate order quantity $\widehat{Q}$ and to the order quantity $\widetilde{Q}$ (resulting from the order rule) currently used by the retailer. The current order quantity is equal to the expected net demand, i.e.,

$$\widetilde{Q} = \mu_P(1 - rk). \tag{13}$$

Recall from Section 2 that $\widetilde{Q}$ is actually an order-up-to level instead of an order quantity, since the preview order on which we focus is the second order. However, as we argued in that section, the preview order is the key order and we consider it to be the only order in this study.

Besides a comparison of Q^*, $\widehat{Q}$, and $\widetilde{Q}$ for all products together, we illustrate their differences for a selection of 9 products. The relevant cost and demand data are given in Table 5.

INSERT TABLE 5 ABOUT HERE

These 9 products were selected to display the effect of an product's profit margin, its expected gross demand and its expected return probability on the order quantities and on the associated expected profits. Products 1-3 in the table are chosen according to decreasing relative profit margin $(p-c)/c$, products 4-6 according to decreasing expected gross demand μ_G, and products 7-9 according to decreasing expected return probability r. Moreover, to isolate the effect of the decreasing parameter, the products are chosen such that all other parameters are approximately constant over every three products for which one of the aforementioned parameters is decreasing.

To display the effect of the shortage cost g (equal for all products), we use three different values: $g = 0$, $g = 10$, and $g = 50$.

7.1 Results

We first compare the order quantities and expected profits for the nine selected products, and will then discuss cumulated results for all products in Set 2. Table 6 displays the results for the selected nine products. Table 7 in Appendix B gives a detailed view of the separate revenues and costs.

INSERT TABLE 6 ABOUT HERE

The first main conclusion is that the current order quantity $\widetilde{Q}$ is too small, often more than 10% smaller than the optimal order quantity Q^*. This holds even if if the shortage cost g is set to 0, though the differences are larger, of course, for $g = 10, 50$. The effect on the expected profit is small (up to 4% decrease) if $g = 0$, considerable (up to 13% decrease) if $g = 10$, and can be very large (up to 78% decrease) if $g = 50$.

The result that $\widetilde{Q} < Q^*$ even for $g = 0$ shows that, by ordering more, the retailer can increase profit (excluding loss of goodwill costs) and reduce the number of lost sales at the same time. The poor performance of $\widetilde{Q}$ determined by (13) can be attributed to its simplicity, and especially to the inability to take into account relevant cost parameters. Besides not considering the shortage cost g, $\widetilde{Q}$ does not differentiate between highly and less profitable products. Clearly, it is better to order relatively more products with a high profit margin. Products 1-3 in Table 6 illustrate this.

Next, we compare $\widehat{Q}$ to Q^* for the selection of 9 products. It turns out that $\widehat{Q}$ is always larger than the optimal order quantity, 13% on average. Recall that $\widehat{Q}$ is based on the assumptions that items can be resold only once and that a fixed percentage of returns is resalable. Recall further that the first assumption leads to an upward bias of the order quantity whereas the second one leads to a downward bias. Apparently, the effect of the 'single resale' assumption is dominant. This also explains why the difference between $\widehat{Q}$ and Q^* is especially large if the return rate is high, since a higher return rate increases the probability that products are resold more than once.

Table 6 further shows that the difference between $\widehat{Q}$ and Q^* is almost constant in g. The difference between the associated expected profits, however, increases considerably in g. The explanation for this result is that, as for the traditional news vendor problem without returns, the expected profit curve $EP(Q)$ is steeper on the right hand side for larger values of g.

We end with a comparison of Q^*, $\widehat{Q}$, and $\widetilde{Q}$ for all products in Set 2 together. Table 8 gives the cumulative numbers for the expected profits and the expected lost sales percentages. We remark that a comparison of the realized profits and the realized lost sales percentages, calculated using the realized demand data in Set 2, produced similar results.

INSERT TABLE 8 ABOUT HERE

Table 8 shows that the percentage of lost sales associated with $\widehat{Q}$ is relatively much smaller than the percentage of lost sales associated with the optimal order quantity Q^*. This is because $\widehat{Q}$ is often much larger than Q^* (see the previous discussion of the results for the selection of 9 products). The overall effect on the profit is small, however, since the profit curve is rather flat around the optimum, especially for small values of g. Ordering $\widehat{Q}$ instead of Q^* leads to a profit reduction of 1% for $g = 0$, 2% for $g = 10$, and 4% for $g = 50$.

The overall performance of the current order quantity is very poor. The associated lost sales percentage is 13%, whereas the optimal lost sales percentage is at most 5% (less if $g > 0$). Using $\widetilde{Q}$ instead of Q^* leads to a profit reduction of 2% for $g = 0$, 5% for $g = 10$, and 38% for $g = 50$.

8 Conclusion

We derived a simple closed-form formula that determines the optimal order quantity Q^* for a single period inventory (newsboy) problem with returns. Using real data from a large catalogue/internet mail order company, Q^* was compared to an approximation $\widehat{Q}$ proposed in a previous study and to the order quantity $\widetilde{Q}$ currently used by the company. It turned out that $\widehat{Q}$ differs more than 10% from Q^* in most cases. The associated profit reduction is generally smaller than 5%, but more than 10% in cases with a high return rate and a high shortage cost. The company's order quantity $\widetilde{Q}$ is far from optimal. It is much smaller (often more than than 20%) than Q^*. Even if the shortage cost is ignored, the company could increase profit by ordering more (while reducing the number of lost sales at the same time).

Due to a lack of data (only forecasts and realizations for multiple items for a single period are available), we had to develop a procedure for estimating the variance of demand. The simple procedure that we developed suffices for the purposes of this paper. It would be interesting, though, to do further statistical research into this forecasting problem and compare different procedures.

In this paper, we focussed on the second of the mail order company's three order moments, the preview. It would be interesting to extend our model with especially an additional in-season emergency replenishment option. Such a model can provide useful insights into the simultaneous determination of the two optimal order quantities and into the profitability of the additional replenishment option.

Acknowledgements:

The research of Dr. Ruud H. Teunter has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

The research presented in this paper is part of the research on re-use in the context of the EU sponsored TMR project REVersed LOGistics (ERB 4061 PL 97-5650) in which the Erasmus University Rotterdam (NL), the Otto-von-Guericke Universität Magdeburg (D), the Eindhoven University of Technology (NL), INSEAD (F), the Aristoteles University of Thessaloniki (GR), and the University of Piraeus (GR) take part.

References

[1] Moutaz Khouja. The Single-period (News-vendor) Problem: Literature Review and Suggestions for Future Research. *Omega, The International Journal of Management Science*, 27:537–553, 1999.

[2] Edward A. Silver, David F. Pyke, and Rein Peterson. *Inventory Management and Production Planning and Scheduling*. John Wiley & Sons, New York, third edition, 1998.

[3] Dimitrios Vlachos and Rommert Dekker. Ordering Policies for Single Period Products with Returns. Technical report, Erasmus University Rotterdam, 2000.

A Mean and variance of net demand

Using

$$\begin{aligned} E[K] &= \sum_{n=0}^{\infty} \Pr[G=n] E[K|G=n] \\ &= \sum_{n=0}^{\infty} \Pr[G=n] rkn \\ &= rkE[G], \end{aligned}$$

$$\begin{aligned} E[K^2] &= \sum_{n=0}^{\infty} \Pr[G=n] E[K^2|G=n] \\ &= \sum_{n=0}^{\infty} \Pr[G=n](V[K|G=n] + (E[K|G=n])^2) \\ &= \sum_{n=0}^{\infty} \Pr[G=n](nrk(1-rk) + (rkn)^2) \\ &= rk(1-rk)E[G] + (rk)^2 E[G^2] \\ &= rk(1-rk)E[G] + (rk)^2(V[G] + (E[G])^2) \\ &= rk(1-rk)E[G] + (rk)^2 V[G] + (rk)^2(E[G])^2, \end{aligned}$$

and

$$\begin{aligned} E[GK] &= \sum_{n=0}^{\infty} \Pr[G=n] E[GK|G=n] \\ &= \sum_{n=0}^{\infty} \Pr[G=n] nE[K|G=n] \\ &= \sum_{n=0}^{\infty} \Pr[G=n] n(rkn) \\ &= rkE[G^2] \\ &= rk(V[G] + (E[G])^2) \\ &= rkV[G] + rk(E[G])^2, \end{aligned}$$

we get

$$
\begin{aligned}
\mu_N &= E[N] \\
&= E[G-K] \\
&= E[G]-E[K] \\
&= E[G]-rkE[G] \\
&= (1-rk)E[G] \\
&= (1-rk)\mu_G
\end{aligned}
$$

and

$$
\begin{aligned}
(\sigma_N)^2 &= V[N] \\
&= V[G-K] \\
&= V[G]+V[K]-2\mathrm{Cov}[G,K] \\
&= V[G]+E[K^2]-(E[K])^2-2(E[GK]-E[G]E[K]) \\
&= V[G]+rk(1-rk)E[G]+(rk)^2V[G]+(rk)^2(E[G])^2-(rk)^2(E[G])^2 \\
&\quad -2\left(rkV[G]+rk(E[G])^2-E[G]rkE[G]\right) \\
&= (1+(rk)^2-2rk)V[G]+rk(1-rk)E[G] \\
&= (1-rk)^2V[G]+rk(1-rk)E[G] \\
&= (1-rk)^2(\sigma_G)^2+rk(1-rk)\mu_G.
\end{aligned}
$$

B Figures

INSERT TABLE 7 ABOUT HERE

G	gross demand
f_G *	density function of gross demand
F_G *	distribution function of gross demand
μ_G	mean of gross demand
σ_G	standard deviation of gross demand
r	expected probability that a sold product is returned
k	expected probability that a returned product is resalable
K	number of resalable returns if all demands are met
N	net demand ($N = G - K$)
f_N *	density function of net demand
F_N *	distribution function of net demand
μ_N	mean of net demand
σ_N	standard deviation of net demand
p	selling price
s	salvage value
c	purchase cost
g	(gross) shortage/loss of goodwill cost
g_N	net shortage cost $g/(1 - rk)$
d	return collection cost
p_G	expected gross revenue $(1 - r)p - rd + r(1 - k)s$
p_N	expected net revenue $p_G/(1 - rk)$
Q	order quantity
$\widetilde{Q}$	order(-up-to) quantity currently used by the retailer
$\widehat{Q}$	order quantity resulting from the approximate analysis
Q^*	order quantity resulting from the exact analysis
$EP(Q)$	expected profit for order quantity Q
$\widehat{EP}(Q)$	approximation of the expected profit for order quantity Q
$\widehat{ES}_G(Q)$	approximation of the expected gross shortage for order quantity Q, i.e. approximation of the expected number of gross demands not met
$ES_N(Q)$	expected net shortage for order quantity Q, i.e. the expected number of net demands not met

*: if demand is approximated by a continuously distributed variable

Table 1: Notations.

	Set 1 (4761 products)	Set 2 (427 products)
demand preview μ_P	4-12203	103-4174
demand realization	1-13195	15-7102
return rate preview r_P	n.a.	36.7%-53.3%
return rate realization	n.a.	18.3%-74.1%
purchase cost c	n.a.	5.25-30.64
return collection cost d	n.a.	4.25
selling price p	n.a.	19.95-99
salvage value s	n.a.	1.58-9.19

Table 2: Data ranges.

		***c* (corresponding b follows from (10))**										
products	$\boldsymbol{\mu_G}$	**1**	**1.1**	**1.2**	**1.3**	**1.4**	**1.5**	**1.6**	**1.7**	**1.8**	**1.9**	**2**
1-779	128-199	66.63	40.07	24.10	14.49	8.72	5.25	3.16	1.90	1.14	0.69	0.41
780-1708	200-299	86.62	49.99	28.85	16.66	9.62	5.55	3.21	1.85	1.07	0.62	0.36
1709-2337	300-399	130.56	72.79	40.58	22.63	12.62	7.04	3.92	2.19	1.22	0.68	0.38
2338-2994	400-599	123.98	66.80	36.00	19.40	10.46	5.64	3.04	1.64	0.88	0.48	0.26
2995-3370	600-799	150.34	78.19	40.67	21.16	11.01	5.73	2.98	1.55	0.81	0.42	0.22
3371-3559	801-999	181.99	92.31	46.82	23.75	12.05	6.11	3.10	1.57	0.80	0.40	0.21
3560-3763	1006-1493	264.93	130.20	63.99	31.45	15.46	7.60	3.74	1.84	0.90	0.44	0.22
3764-3853	1500-1997	291.73	138.30	65.57	31.09	14.74	6.99	3.32	1.57	0.75	0.35	0.17
3854-1913	2017-2986	562.73	258.60	118.86	54.63	25.12	11.55	5.31	2.44	1.12	0.52	0.24
3914-3937	3200-10445	593.08	258.43	112.70	49.18	21.48	9.39	4.10	1.80	0.79	0.34	0.15

Table 3: Average value of $(G - \mu_G)^2/(\mu_G)^c$ for different values of c and different ranges of μ_G.

Product	Previews		Gross demand		Net demand	
	μ_P	r_P	μ_G	σ_G	μ_N	σ_N
I	150	38	129	84	82	54
II	808	39	692	352	434	221
III	4174	39	3573	1420	2249	895

Table 4: Three products from Set 1 for which Normality of net demand is tested.

product	c	p	$(p-c)/c$	s	$r = r_P$	μ_P	μ_G	σ_G	μ_N	σ_N
1	7.56	35.00	**3.63**	2.27	0.37	545	466	251	301	163
2	14.02	49.95	**2.56**	4.21	0.37	545	466	251	301	163
3	16.35	38.85	**1.38**	4.91	0.37	545	466	251	301	163
4	30.64	89.95	1.94	9.19	0.39	3451	**2954**	1208	1860	761
5	13.66	39.95	1.92	4.10	0.40	1253	**1072**	511	662	315
6	13.66	39.95	1.92	4.10	0.41	478	**409**	225	250	138
7	14.85	49.95	2.36	4.46	**0.53**	572	490	262	242	130
8	17.28	59.95	2.47	5.18	**0.44**	566	484	260	280	151
9	8.75	29.90	2.42	2.63	**0.37**	599	513	273	331	177

Table 5: Data for the selected group of 9 products from Set 2. Note that $k = 0.95$ and $d = 4.25$ for all products.

product	Q^*	$\widehat{Q}$	$\widetilde{Q}$	$EP(Q^*)$	$EP(\widehat{Q})$	$EP(\widetilde{Q})$
	order quantity			expected profit		
g = 0						
1	450	496 (+10%)	352 (-22%)	5979	5932 (-1%)	5715 (-4%)
2	419	459 (+10%)	352 (-16%)	7582	7522 (-1%)	7388 (-3%)
3	353	378 (+7%)	352 (-0%)	3864	3842 (-1%)	3864 (-0%)
4	2295	2546 (+11%)	2172 (-5%)	81245	80254 (-1%)	80985 (-0%)
5	828	917 (+11%)	773 (-7%)	11296	11164 (-1%)	11242 (-0%)
6	323	356 (+10%)	293 (-9%)	4047	4005 (-1%)	4009 (-1%)
7	321	386 (+20%)	282 (-12%)	5133	4951 (-4%)	5054 (-2%)
8	385	438 (+14%)	327 (-15%)	8119	7984 (-2%)	7937 (-2%)
9	448	490 (+9%)	387 (-14%)	4570	4534 (-1%)	4483 (-2%)
g = 10						
1	494	549 (+11%)	352 (-29%)	5791	5719 (-1%)	5055 (-13%)
2	456	503 (+10%)	352 (-23%)	7302	7209 (-1%)	6729 (-8%)
3	412	450 (+9%)	352 (-15%)	3374	3312 (-2%)	3204 (-5%)
4	2411	2691 (+12%)	2172 (-10%)	79368	78053 (-2%)	78244 (-1%)
5	929	1045 (+12%)	773 (-17%)	10561	10304 (-2%)	9976 (-6%)
6	367	413 (+13%)	293 (-20%)	3722	3632 (-2%)	3412 (-8%)
7	362	449 (+24%)	282 (-22%)	4805	4451 (-7%)	4363 (-9%)
8	418	482 (+15%)	327 (-22%)	7809	7599 (-3%)	7251 (-7%)
9	511	565 (+11%)	387 (-24%)	4270	4198 (-2%)	3769 (-12%)
g = 10						
1	569	638 (+12%)	352 (-38%)	5454	5328 (-2%)	2419 (-56%)
2	527	589 (+12%)	352 (-33%)	6728	6552 (-3%)	4092 (-39%)
3	505	562 (+11%)	352 (-30%)	2530	2361 (-7%)	568 (-78%)
4	2687	3031 (+13%)	2172 (-19%)	74687	72453 (-3%)	67283 (-10%)
5	1096	1251 (+14%)	773 (-29%)	9265	8735 (-6%)	4912 (-47%)
6	441	503 (+14%)	293 (-34%)	3153	2951 (-6%)	1025 (-67%)
7	430	549 (+28%)	282 (-34%)	4231	3544 (-16%)	1598 (-62%)
8	484	567 (+17%)	327 (-32%)	7159	6766 (-5%)	4510 (-37%)
9	605	678 (+12%)	387 (-36%)	3789	3643 (-4%)	913 (-76%)

Table 6: Results for a selection of 9 products from Set 2. The percentual deviations are relative to the optimal order quantity Q^* and to the associated optimal profit $EP(Q^*)$.

pro-duct			$\widetilde{Q}$					$\widehat{Q}$					$Q^\star$		
	p	s	c	d	g	p	s	c	d	g	p	s	c	d	g
$g = 0$															
1	8780	229	-2660	-634	-0	9917	483	-3751	-716	-0	9685	392	-3399	-700	-0
2	12530	425	-4933	-634	-0	13904	762	-6440	-704	-0	13517	625	-5875	-684	-0
3	9761	497	-5762	-633	-0	10095	579	-6178	-654	-0	9774	499	-5776	-633	-0
4	147048	4941	-66563	-4442	-0	155450	7518	-78018	- 4696	-0	150378	5727	-70317	-4543	-0
5	22572	854	-10563	-1621	-0	24144	1283	-12529	-1734	- 0	23272	1007	-11312	-1671	-0
6	8261	352	-3999	-606	-0	8987	536	-4859	-659	-0	8655	435	-4409	-635	-0
7	9818	383	-4193	-953	-0	11016	739	-5734	-1070	-0	10414	504	-4774	-1011	-0
8	13882	496	-5655	-786	-0	15503	930	-7571	-878	-0	14904	704	-6645	-844	-0
9	8279	289	-3384	-700	-0	9100	487	-4283	-769	-0	8834	400	-3917	-747	-0
$g = 10$															
1	8780	229	-2660	-634	-659	10073	593	-4150	-728	-70	9906	478	-3732	-716	-146
2	12530	425	-4933	-634	-659	14192	924	-7059	-718	- 129	13874	749	-6388	-702	-230
3	9761	497	-5762	-633	-659	10768	850	-7362	-698	-247	10458	700	-6733	-678	-373
4	147048	4941	-66563	-4442	-2740	157466	8644	-82459	- 4757	-842	152992	6525	-73871	-4622	-1657
5	22572	854	-10563	-1621	-1266	24926	1728	-14281	- 1790	-279	24240	1323	-12695	-1741	-566
6	8261	352	-3999	-606	-597	9357	732	-5637	-686	-134	9083	575	-5020	-666	-250
7	9818	383	-4193	-953	-691	11286	997	-6674	-1096	-62	10843	648	-5382	-1053	-252
8	13882	496	-5655	-786	-685	15800	1132	-8329	-894	- 110	15309	841	-7217	-867	-257
9	8279	289	-3384	-700	-714	9397	661	-4947	-795	-118	9203	533	-4467	-778	-221
$g = 50$															
1	8780	229	-2660	-634	-3296	10188	788	-4824	-736	-88	10112	636	-4303	-730	-260
2	12530	425	-4933	-634	-3296	14473	1258	-8252	-733	- 195	14299	1014	-7390	-724	-471
3	9761	497	-5762	-633	-3296	11226	1338	-9181	-727	- 295	11057	1080	-8250	-716	-640
4	147048	4941	-66563	-4442	-13702	160307	11476	-92868	-4843	-1620	157421	8615	-82345	-4755	-4249
5	22572	854	-10563	-1621	-6330	25444	2518	-17090	- 1827	-309	25118	1916	-14973	-1804	-992
6	8261	352	-3999	-606	-2985	9602	1078	-6877	-704	- 149	9467	835	-6021	-694	-435
7	9818	383	-4193	-953	-3456	11411	1428	-8146	-1108	- 41	11226	914	-6380	-1090	-438
8	13882	496	-5655	-786	-3427	16070	1549	-9801	-910	- 144	15808	1139	-8355	-895	-537
9	8279	289	-3384	-700	-3571	9569	941	-5929	-809	-128	9484	757	-5295	-802	-355

Table 7: Detailed revenue/cost results for a selection of 9 products from Set 2. Separate returns and costs are shown for all three order quantities, $\widetilde{Q}$, $\widehat{Q}$ and $Q^\star$.

g	$\boldsymbol{EP(Q^*)}$	$\boldsymbol{EP(\widehat{Q})}$	$\boldsymbol{EP(\widetilde{Q})}$	$\boldsymbol{LS(Q^*)}$	$\boldsymbol{LS(\widehat{Q})}$	$\boldsymbol{LS(\widetilde{Q})}$
	expected profit			**lost sales percentage**		
0	3190340	3156470 (-1%)	3127520 (-2%)	8%	5%	13%
10	3058070	3004680 (-2%)	2850970 (-5%)	5%	3%	13%
50	2797180	2696130 (-4%)	1744770 (-38%)	2%	1%	13%

Table 8: Cumulated results for all products in Set 2. The percentual deviations for the expected profits are relative to $EP(Q^*)$. The lost sales percentage $LS(Q)$ is the fraction of gross demands that are lost (one minus the 'fill rate').

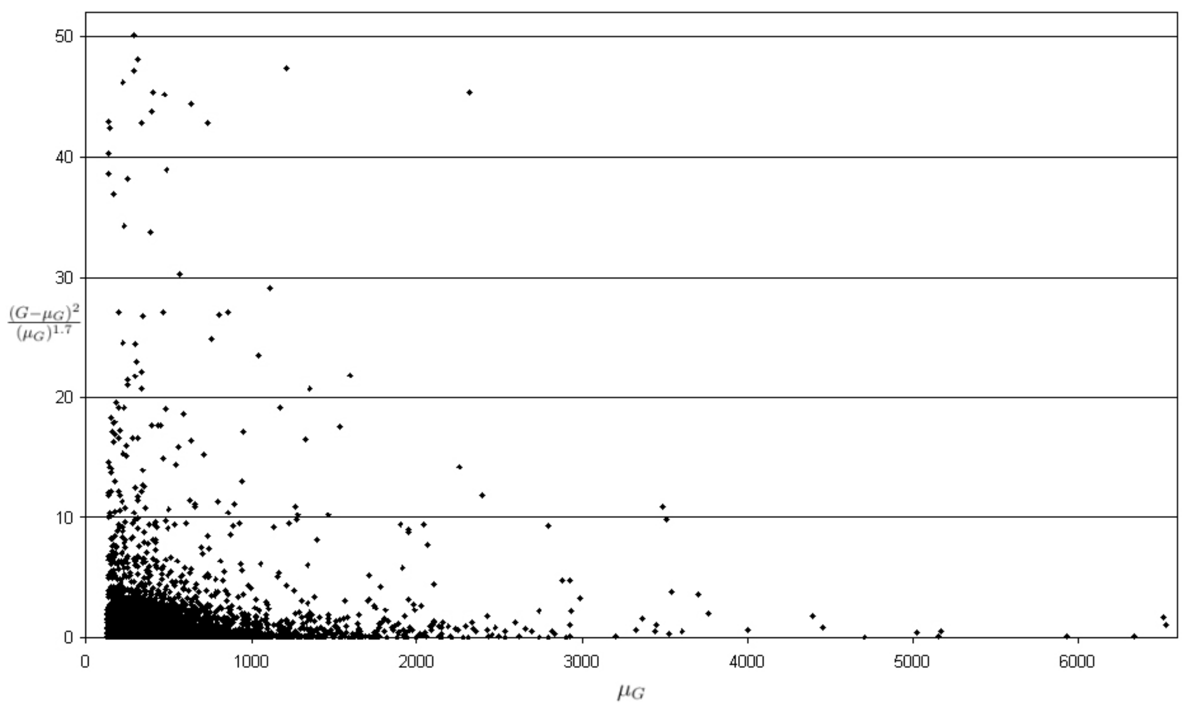

Figure 1: Scatterplot of $(G - \mu_G)^2/(\mu_G)^{1.7}$ against μ_G. To get a clearer picture, we left out 3 clear outliers and one product with an extremely high value of μ_G.

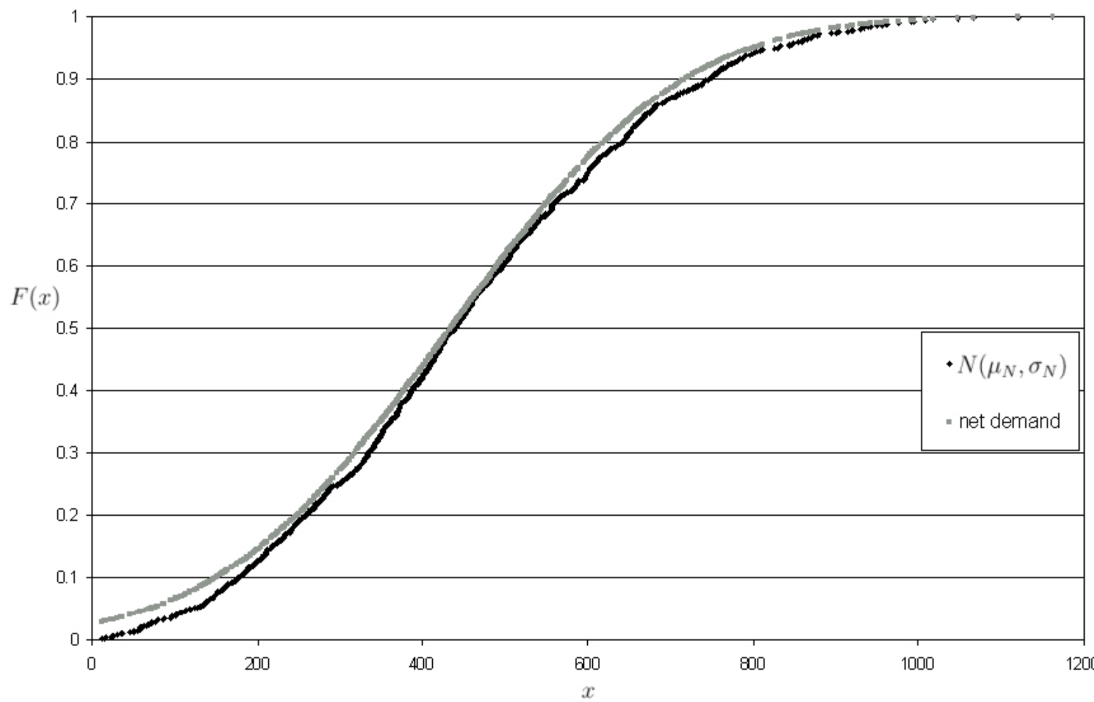

Figure 2a: Scatter-plot for product 1 (see table 4) of the cumulative distribution function of net demand computed from 1000 random drawings against the Normal cdf with mean μ_N and standard deviation σ_N.

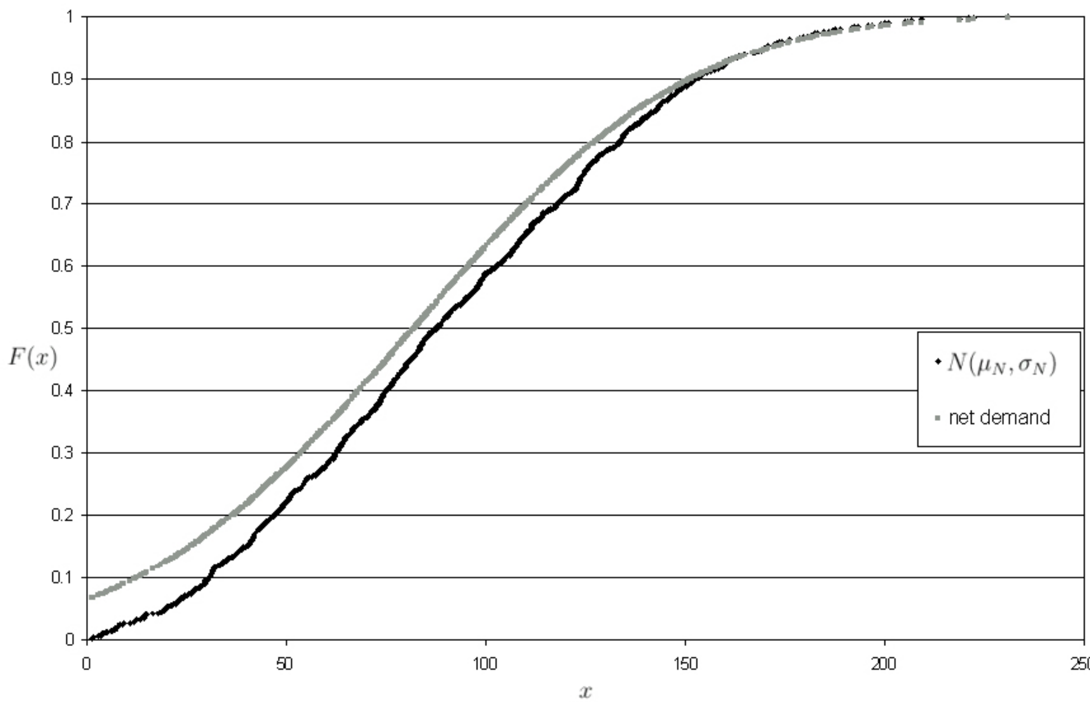

Figure 2b: Scatter-plot for product 2 (see table 4) of the cumulative distribution function of net demand computed from 1000 random drawings against the Normal cdf with mean μ_N and standard deviation σ_N.

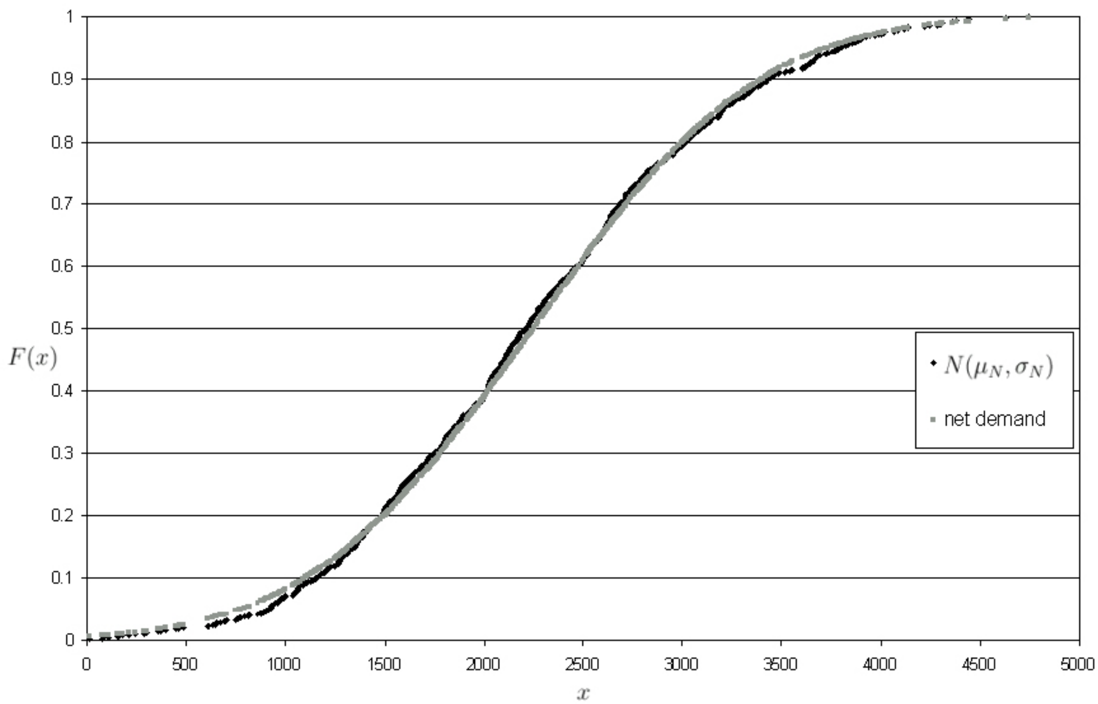

Figure 2c: Scatter-plot for product 3 (see table 4) of the cumulative distribution function of net demand computed from 1000 random drawings against the Normal cdf with mean μ_N and standard deviation σ_N.