Inventory Management with product returns: the value of information

Marisa P. de Brito and E. A. van der Laan

ERIM REPORT SERIES *RESEARCH IN MANAGEMENT*	
ERIM Report Series reference number	ERS-2003-060-LIS
Publication	2003
Number of pages	34
Email address corresponding author	debrito@few.eur.nl
Address	Erasmus Research Institute of Management (ERIM) Rotterdam School of Economics / Faculteit Economische Wetenschappen Rotterdam School of Management / Faculteit Bedrijfskunde Erasmus Universiteit Rotterdam P.O. Box 1738 3000 DR Rotterdam, The Netherlands Phone: +31 10 408 1182 Fax: +31 10 408 9640 Email: info@erim.eur.nl Internet: www.erim.eur.nl

Bibliographic data and classifications of all the ERIM reports are also available on the ERIM website:
http://www.erim.eur.nl/

ERASMUS RESEARCH INSTITUTE OF MANAGEMENT

REPORT SERIES
RESEARCH IN MANAGEMENT

BIBLIOGRAPHIC DATA AND CLASSIFICATIONS		
Abstract	This paper evaluates the impact of misinformation for inventory systems with product returns. If one could exactly know how much is going to be returned and when, one would certainly benefit from incorporating this perfect information a priori in the management of production, inventory, and distribution. In practice, one has to attempt to forecast the timing and the amount of product returns, by hypothesizing about the return flow properties. To do so, historic data on demand and returns can be used. The available literature on information and inventory management with product returns commonly 1) assumes known return probabilities; or 2) considers specific cases where the most informed method does not necessarily lead to the best performance, investigating the impact on inventory related costs.	
Library of Congress Classification (LCC)	5001-6182	Business
	5201-5982	Business Science
	HD 40	Inventory policy
Journal of Economic Literature (JEL)	M	Business Administration and Business Economics
	M 11 R 4	Production Management Transportation Systems
	M 11	Production managemt
European Business Schools Library Group (EBSLG)	85 A	Business General
	260 K 240 B	Logistics Information Systems Management
	260 M	Inventory and materials management
Gemeenschappelijke Onderwerpsontsluiting (GOO)		
Classification GOO	85.00	Bedrijfskunde, Organisatiekunde: algemeen
	85.34 85.20	Logistiek management Bestuurlijke informatie, informatieverzorging
	85.35	Productiemanagement
Keywords GOO	Bedrijfskunde / Bedrijfseconomie	
	Bedrijfsprocessen, logistiek, management informatiesystemen	
	Voorraadbeheer, informatie management, forecasting	
Free keywords	Product returns, inventory management, information management, forecasting	

Inventory Management with product returns: the value of information

Marisa P. de Brito
Erasmus University Rotterdam,
Econometric Institute, Room H10-14,
PO Box 1738, 3000 DR Rotterdam,
The Netherlands
debrito@few.eur.nl

E. A. van der Laan
Erasmus University Rotterdam,
Faculty of Business Administration
elaan@fbk.eur.nl

Version July-2003

Abstract

This paper evaluates the impact of misinformation for inventory systems with product returns. If one could exactly know how much is going to be returned and when, one would certainly benefit from incorporating this perfect information a priori in the management of production, inventory, and distribution. In practice, one has to attempt to forecast the timing and the amount of product returns, by hypothesizing about the return flow properties. To do so, historic data on demand and returns can be used. The available literature on information and inventory management with product returns commonly 1) assumes known return probabilities; or 2) considers specific cases where the most-informed method leads to the best forecast. This paper identifies situations in which the most informed method does not necessarily lead to the best performance, investigating the impact on inventory related costs.

Keywords: Product returns, inventory management, information management, forecasting.

1 Introduction

Products and packaging return into the supply chain for a diversity of reasons (see De Brito and Dekker, 2003). For instance, beverage containers are returned by the consumer to the retailer against reimbursement, single-use photo cameras are turned in for the film to be developed, and now millions of products purchased through mail-order-companies, e-tailers and other distant sellers, are being returned every day.

There is a lot of money involved with the handling of product returns (Rogers and Tibben-Lembke, 1999). One of the difficulties in handling returns efficiently is that return flows are often characterized by a considerable uncertainty essentially regarding time and quantity. If one could exactly know how much is going to be returned and when, one would certainly benefit from incorporating this perfect information a priori in the management of production, inventory, and distribution. In many cases this is far from feasible (see Trebilcock, 2002). Nevertheless, one may attempt to forecast the timing and the amount of product returns. To do so, one has to hypothesize about the return flow properties based on historic demand and return data.

There are not many papers that investigate the impact of information on inventory management with product returns. The ones that do so assume known return probabilities or consider specific cases where the most-informed method leads to the best forecast. However, in environments where data is scarce or unreliable, or in environments that are volatile, information may be misleading. This study reports on the impact that such (mis)information has on inventory management. This paper identifies situations in which the most informed method does not necessarily lead to the best performance. Furthermore, the impact on inventory related costs of having inaccurate estimates of the time-to-return distribution is investigated for a wide range of parameters. In addition, the implications for the practice of inventory management with returns will be discussed.

The remainder of the paper is structured as follows. First, Section 2 illustrates that estimates of the return rate can be massively erroneous in practice. Next, Section 3 reviews the literature on the topic. Subsequently Section 4 elaborates on procedures to estimate the net demand. By means of an exact analysis, Section 5 identifies situations in which more information does not necessarily lead to better performance. Section 6 then presents a simulation study to quantify the impact of misinformation with respect to inventory related costs. In Section 7, the managerial implications are discussed and Section 8 concludes by giving recommendations for future research.

2 (Mis)information in practice

This paper reports on the impact of (mis)information in an inventory system with product returns. Misinformation means here that the properties we assume for the *future* return flow do not correspond to the actual properties of the *future* return flow. Further attention is paid next on how misinformation can occur in a given system.

Figure 1 illustrates how the "information" arrives to the decision-maker. A tool transforms data into "information" (e.g. system parameters), with which the decision-maker controls the system. On the one hand, data may be insufficient, abundant, ambiguous or conflicting (see Zimmerman, 1999) causing misinformation as input. On the other hand, the tool may be imperfect creating some sort of distortion. The impact of misinformation on inventory systems has been studied for long. Bodt and Van Wassenhove (1983) explained how firms could see the advantages of using Material Requirements Planning (MRP) being drastically reduced due to misinformation. Gardner (1999) showed how important is to choose an adequate forecasting tool while investing in customer service levels. Some misinformation mechanisms are well-known in inventory management, like the bullwhip effect, as introduced by Lee et al. (1997). Misinformation is a general phenomenon, not specific for systems with product returns. Yet, academics and practitioners agree that managing systems with product returns is very demanding with respect to neat information, which is sparsely available (see Gooley, 1998; Guide, 2000; De Brito et al.,2003; Kokkinaki et al., 2003).

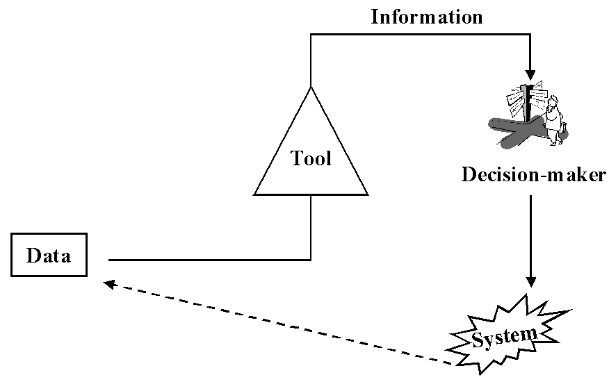


Figure 1: Relation between data and decision-making

In this paper, the source or reason of misinformation is not a matter of concern. The objective of this study is to evaluate the impact of misinformation on the system, given that misinformation occurs. Next a real-life example illustrates that misinformation on product returns does happen in practice. The company's name is here disguised as MOC.

MOC is a large Western European mail-order-company. Well before the start of the season, the company has to decide on what to offer for sale in the upcoming catalogue. A great part of the total stock to cover all the season has to be ordered many weeks in advance because production lead times are long. This is many times the case in Europe or the U.S. where manufacturing comes from Asia (see Mostard and Teunter, 2002). The order decisions take into account the expected future demand and returns for each line of products. Since there is however no data available on the sales of the coming season, the forecast is not much more than a rough estimate. In particular, a mail-order-company in the Netherlands uses estimates that on average are more than 20% off the real return rate and sometimes even more than 80%. The above illustrates that large mis-estimation of the return rate does occur in practice. This paper analyzes its impact on inventory management.

3 Literature Review

The literature dealing with product returns has been growing fast in the last years. This literature falls in the general umbrella of closed-loop supply chain management. For contextual contents, see Rogers and Tibben-Lembke (1999); Guide and Van Wassenhove (2003); Dekker et al. (2003). Inventory Management has received chief attention from the beginning until now, which has brought about a richness of studies (see e.g. the following recent ones: Van der Laan et al., 1999; Inderfurth et al., 2001; Fleischmann et al., 2002). In spite of these and many other contributions, there are few articles that simultaneously consider the forecasting of product returns, inventory management and information issues. Below follows a review on specific literature contributing to close the aforementioned gap.

Goh & Varaprasad (1986) developed a methodology to compute life-cycle parameters of returned containers by employing data on demand and returns. The authors claim that a careful estimation of these parameters aids to effective inventory management. They apply their method to data on soft drinks as Coca-Cola and Fanta from the Malaysia and Singapore markets. The approach requires a time series of aggregated demand and another one of aggregated return data. They used a 50-point time series of monthly data. The precision of the method depends on how accurate the estimation of the return distribution is. This estimation has the Box-Jenkins time series techniques as basis. The authors call the attention of the reader to the fact that a data set with less than 50 points is too short to employ the methodology. Large time series should also be avoided. Therefore, they recommend a time series of 50 points, or a 4-year period of data coming from a stable market environment. Our research contributes to contrasting situations: with imperfect or misleading data.

Kelle & Silver (1989) proposed four forecasting procedures of net demand during lead time in the case of reusable containers. Every procedure has a different level of information requirement. The least-informed method uses the expectation and the variance of the net demand together with the probability of return. The most-informed one calls for individual tracking and tracing of containers. The authors evaluate the forecasting methods taking the most-informed method as a benchmark. The analysis, however, applies only to the case of perfect information on the return parameters. This paper employs the methods proposed by Kelle and Silver (1989) but it elaborates on the potential impact of (mis)information.

Toktay et al. (2000) consider the real case of new circuit boards for Kodak's single use remanufactured camera. The goal is to have an ordering policy that minimizes the procurement, inventory holding and lost sales costs. A six-node closed queueing network is employed to represent Kodak's supply chain. Accordingly, the returns of cameras depend on past sales by a return probability and an exponential time lag distribution. Some of the procedures of Kelle and Silver are used to predict the unobservable inventory at the customer-use network node. The authors compare several forecasting methodologies with different levels of information. This paper provides an exact and more general analysis of the impact of (mis)information. Furthermore, the four methods of Kelle and Silver are compared numerically for a wide range of parameter values with respect to misinformation. Therefore, besides an intra-method comparison, here an inter-method analysis is also presented.

Marx-Gómez et al. (2002) consider a case of photocopiers that may return to the producer after being used. The authors put forward a method to forecast the number and time of returned photocopiers. Firstly, data is generated according to two scenarios: successful vs. not so successful return incentives. These scenarios, together with expert knowledge, constitute the basis for developing a set of forecasting rules. The expert evaluates factors, such as demand and life cycle parameters. An extended approach of the model is suggested by allowing a follow-up period to self-learn the rules. This neuro-fuzzy process calls for data on demands and returns. Besides this, one should notice that the method of Marx-Gómez et al. depends on a priori knowledge on product returns, which have been acquired by the producer.

Overall one can conclude that the proposed procedures are very demanding with respect to reliable data. This paper takes the more realistic perspective of misinformation and analyzes its impact.

Table 1: Notation.

t,	the current time period; $t = 0, 1, 2, \ldots$
L,	order leadtime
$D_L(t)$,	r.v. representing the demand during the interval $[t+1, t+2, \ldots, t+L]$
$R_L(t)$,	r.v. representing the returns during the interval $[t+1, t+2, \ldots, t+L]$
$ND_L(t)$,	the net demand during the leadtime (equal to $D_L(t) - R_L(t)$).
p_j,	the probability of an item returning after exactly $j = 1, \ldots, n$ periods, n being the largest j for which the return probability is larger than zero.
p,	the probability of an item ever being returned, i.e. $\sum_{j=1}^{n} p_j$.
u_i,	demanded amount during period i and $i \leq t$.
y_i,	the total amount returned in each previous period $i \leq t$.
S,	base-stock level.
k,	safety factor

4 Forecasting Methods

The main notation used in the paper is as in Table 6.1.

Consider a single product, single echelon, periodic review inventory system. Each individual demand returns with probability p according to some return distribution. We assume that this process does not change in time (stationary case). Returns are immediately serviceable. The order lead time is a fixed constant L. Demands that cannot be satisfied immediately are fully backordered. Following the methodology of Kelle and Silver (1989), for each time period t a base-stock policy is applied. If the net lead time demand, $ND_L(t) = D_L(t) - R_L(t)$, follows a normal distribution the optimal base stock level is given as

$$S = \mathrm{E}[ND_L(t)] + k \cdot \sqrt{\mathrm{Var}[ND_L(t)]}, \tag{1}$$

with $E[ND_L(t)]$ and $\mathrm{Var}[ND_L(t)]$ being the expectation and variance of the net demand during lead time. The safety factor k is determined according to some desired performance level (see e.g. Silver et al., 1998).

In order to estimate $\mathrm{E}(ND_L(t))$ and $\mathrm{Var}[ND_L(t)]$, the four methods first put forward by Kelle and Silver (1989) are used. Apart from the expectation and variance of the demand during lead time ($\mathrm{E}[D_L(t)]$ and $\mathrm{Var}[D_L(t)]$ respectively), each method requires a different level of information for estimating the lead time returns. Below, in increasing order of information needs, the four methods, denoted A–D, are listed (see Appendix 1 for details).

Method A - Average behavior
This method requires the following information:

- p, the overall return probability, i.e., the probability that a product is being returned eventually.

This method is an approximation in the sense that all the returns during the lead time are assumed to be perfectly correlated with the demand during that same lead time and independent of previous demands (see Figure 2). Returns during the lead time are estimated through the overall return probability (it is like if a product is to return, it returns instantaneously). Then, demand and returns during the lead time are simply netted. No historical information is used with respect to demands and returns, so in a static environment the resulting base stock level is constant in time.

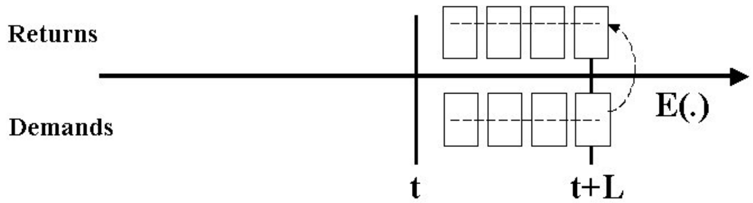


Figure 2: Method A (it uses the overall return probability to compute the expected returns coming from demand during the lead time

Method B - Return distribution

Suppose that we are at the end of period t. This method requires information on previous demand per period and the knowledge of the return distribution as follows:

- u_i, demanded amount during period $i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.

This method makes use of the return probabilities to determine the number and moment of the returns during the order lead time (see Figure 3.

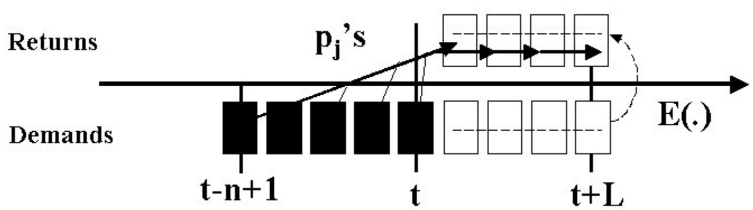


Figure 3: Method B (it uses past demand and return probabilities $p_j's$ to compute returns coming from past demand and it employees expected values for returns coming from demand during the lead time)

Method C - Return distribution & return information per period

Suppose that we are at the end of period t. In addition to the requirements of method B, this method makes use of observed data on aggregated returns:

- u_i, demanded amount during period $i, i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.
- y_i, the total amount of returned products in each period $i, i \leq t$.

This method updates the number of items that have returned, which are observed, and takes this into account to compute future returns (see Figure 4. Thus, Method C aims at improving Method B by taking into account the correlations between the observed aggregated returns in recent periods and the future lead time returns. An analytical method, however, is not available, so Kelle and Silver (1989) developed an approximation (see Appendix 1) that is accurate as long as the purchased amount is relatively large and the return probabilities are positive for several periods (in practice $n \geq 4$).

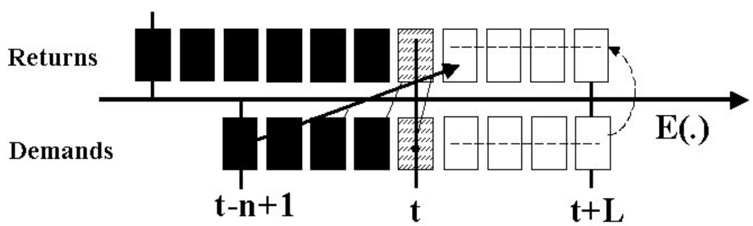


Figure 4: Method C (it uses information on past demand and past returns to compute the returns coming from past demand; it uses information on current demand to compute returns coming from period t; and it uses expected value to compute returns coming from demand during the lead time)

Method D - Return distribution & tracked individual returns

Let t be the last observed period. Besides the requirements of method B this method requires to track back in what period each individual return has been sold:

- u_i, demanded amount during period $i, i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.
- Z_i^t, the observed total number of product returns from each past purchase $u_i, i < t$.

To employ this method, one has to be able to trace back the returns, i.e. to know from which order period each specific return is coming from (see Figure 5.

Given perfect information this method makes optimal use of all relevant information.

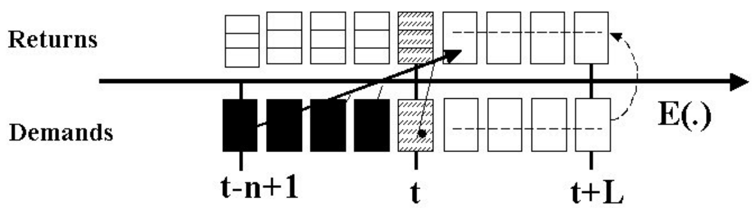


Figure 5: Method D (it uses information on past demand and observed returns from each past demand to compute the returns coming from past demand; it uses information on current demand to compute returns coming from period t; and it uses expected value to compute returns coming from demand during the lead time)

Summarizing, all four methods for estimating the net demand during the order lead time, make use of the expectation and variance of demand. Additionally, each method has different requirements with respect to product return information. Method A is the least demanding: only an estimate of the return rate is needed. Apart from the return rate, Method B requires the return time distribution. On top of that, Method C also needs a record of the aggregated returns per period. Finally, to employ Method D one needs to invest in a system that allows to scan individual returns and track them back to the period in which they were originally sold.

Given perfect information one expects that the method that uses more information outperforms the methods that use less. The remainder of the paper investigates how the various methods perform in presence of misinformation. First, Section 5, identifies situations in which Method B may outperform the most informed method, Method D. Then Section 6 compares Methods A–D with respect to inventory related costs and confirm the findings of Section 5 by means of a simulation study.

5 Forecasting Performance

This section analyzes the relative performance of Methods B and D given misinformation. Methods A and C contain approximations which makes them less interesting for an exact analysis. Besides that, Method A is a rather naive forecasting method, which is not expected to perform very well in general (this will be confirmed in Section 6) and the performance of Method C tends to be very close to that of Method D (Kelle & Silver, 1989).

Since, given perfect estimation, Method D is expected to lead to the best forecast of the expected lead time net demand, we use Method D *given perfect information* as a benchmark for our study. Given *imperfect* estimation, both Methods B and D will do worse than the benchmark, but it is important to know whether there are situations in which Method B outperforms Method D.

Therefore, with respect to the expected lead time net demand we would like to compare

$$\left|\widehat{\mathrm{E}}_B[ND_L(t)] - \mathrm{E}_D[ND_L(t)]\right| \, and \left|\widehat{\mathrm{E}}_D[ND_L(t)] - \mathrm{E}_D[ND_L(t)]\right|$$

Here '$\widehat{\ }$' denotes a forecast, based on forecasts $\{\widehat{p_i}\}$ of $\{p_i\}$. However, the above expression depends on observations Z_i of past product returns coming from demand u_i. Accordingly, we take the conditional expectation with respect to the return process given the history of the demand process, $\mathcal{E}_{R|D}$:

$$F_{\{E\}} = \mathcal{E}_{R|D}\left\{\left|\widehat{\mathrm{E}}_B[ND_L(t)] - \mathrm{E}_D[ND_L(t)]\right| - \left|\widehat{\mathrm{E}}_D[ND_L(t)] - \mathrm{E}_D[ND_L(t)]\right|\right\} \tag{2}$$

Note that if $F_{\{E\}} < 0$ then, on average, Method B outperforms Method D with respect to the expected net leadtime demand and *vice versa* if $F_{\{E\}} > 0$. Similarly, the performance measure with respect to the variance of the lead time net demand is

$$F_{\{V\}} = \mathcal{E}_{R|D}\left\{\left|\widehat{V}(<t)_B - V(<t)_D\right| - \left|\widehat{V}(<t)_D - V(<t)_D\right|\right\} \tag{3}$$

To make the analysis more readable we define $\pi_i = \sum_{j=1}^{t-i} p_j$
and we write '$\sum$' for '$\sum_{i=t-n+1}^{t-1}$'. In section 5.1 we analyze the expectation of the net demand during leadtime and in section 5.2 its variance.

5.1 Analysis regarding the expectation of lead time net demand

Define

$$E_{\{BD\}} = \widehat{\mathrm{E}}_B[ND_L(t)] - \mathrm{E}_D[ND_L(t)],$$
$$E_{\{DD\}} = \widehat{\mathrm{E}}_D[ND_L(t)] - \mathrm{E}_D[ND_L(t)]$$

Obviously, $E_{\{BD\}} = 0$ implies that, given misinformation, Method B performs at least as good as Method D and $E_{\{DD\}} = 0$ implies that, given misinformation, Method D performs as good as the benchmark and therefore performs at least Method B will not have a better performance. Conditioning on the signs of $E_{\{BD\}}$ and $E_{\{DD\}}$ we have 4 remaining cases, EXP1–EXP4, to analyze:

Case EXP1: $E_{\{BD\}} > 0$ and $E_{\{DD\}} > 0$

Case EXP2: $E_{\{BD\}} < 0$ and $E_{\{DD\}} < 0$

Case EXP3: $E_{\{BD\}} < 0$ and $E_{\{DD\}} > 0$

Case EXP4: $E_{\{BD\}} > 0$ and $E_{\{DD\}} < 0$

From the analysis of EXP1–EXP4 in Appendix 2 we conclude that, on average, Method B results in a better forecast of the expected lead time net demand in case we consistently under- or overestimate the return probabilities $pj's$.

From cases EXP1–EXP4 it is also clear that the difference between the methods increases as the R_i get bigger. In other words, for higher return rates the differences between the methods are also larger (with respect to the forecasts of the expected lead time net demand).

5.2 Analysis regarding the variance of lead time net demand

Using expressions (13) and (24) in Appendix 1 we define

$$\begin{aligned} V_{\{BD\}} &= \widehat{\mathrm{Var}}_B[ND_L(t)] - \mathrm{Var}_D[ND_L(t)] \\ &= \sum \left[u_i \widehat{R}_i(1-\widehat{R}_i) - (u_i - Z_i)Q_i(1-Q_i)\right], \\ V_{\{DD\}} &= \widehat{\mathrm{Var}}_D[ND_L(t)] - \mathrm{Var}_D[ND_L(t)] \\ &= \sum (u_i - Z_i)\left[\widehat{Q}_i(1-\widehat{Q}_i) - Q_i(1-Q_i)\right]. \end{aligned}$$

We attempted an analysis of the performance measure with respect to the variance, as defined in (3), similarly to the analysis pursued for the expected value. However, we did not obtain straightforward inequalities (see De Brito and van der Laan, 2002). It is very likely that with respect to the variance there are situations for which Method B outperforms D, but also situations for which D outperforms B. We enhance the analysis by looking again at the special case that we misestimate the overall return probability, i.e. $\widehat{p} = a \cdot p$ for some $a : 0 \leq a \leq 1/p$, while the shape of the return distribution remains intact. We can write $\widehat{R}_i = aR_i$ and $\widehat{Q}_i = aR_i/(1-a\pi_i)$. Still it is difficult to analyze $F_{\{V\}}$ directly because of the summations. So, instead we analyze the individual *coefficients* of u_i in $V_{\{BD\}}$, $V_{\{DD\}}$,and $F_{\{V\}}$. Denote these by $\nu_{\{BD\},i}$, $\nu_{\{DD\},i}$, and ϕ_i. If we define $Z_i = \beta_i u_i$, $0 \leq \beta \leq 1$ we can write

$$\nu_{\{BD\},i} = R_i\left[a(1-aR_i) - \left(1 - \frac{R_i}{1-\pi_i}\right)\left(\frac{1-\beta_i}{1-\pi_i}\right)\right]$$

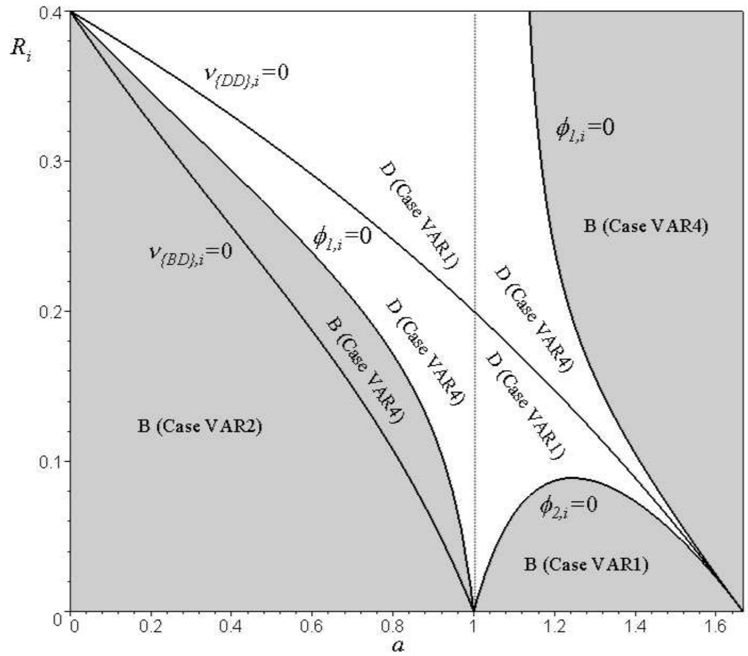
$$\nu_{\{DD\},i} = (1-\beta_i)R_i\left[\frac{a(1-aR_i/(1-a\pi_i))}{1-a\pi_i} - \frac{1-R_i/(1-\pi_i)}{1-\pi_i}\right]$$

$$\phi_i = \begin{cases} \phi_{1,i} = & aR_i\left[(1-aR_i) - \left(1-\frac{aR_i}{1-a\pi_i}\right)\left(\frac{1-\pi_i}{1-a\pi_i}\right)\right] \qquad \text{(VAR1)} \\ \phi_{2,i} = & -\phi_{1,i} \qquad \text{(VAR2)} \\ \phi_{3,i} = & R_i\left[2\left(1-\frac{R_i}{1-\pi_i}\right) - a\left(1-\frac{aR_i}{1-a\pi_i}\right)\left(\frac{1-\pi_i}{1-a\pi_i}\right) - a(1-aR_i)\right] \qquad \text{(VAR3)} \\ \phi_{4,i} = & -\phi_{3,i} \qquad \text{(VAR4)} \end{cases}$$

Figure 6 gives an example of how the preference regions are constructed through $\nu_{\{BD\},i}$, $\nu_{\{DD\},i}$, and ϕ_i in terms of a and R_i. Note that the boundary between the preference regions of Method B and D only depends on ϕ_i (the coefficient of $F_{\{V\}}$) and not on the coefficients of $V_{\{BD\}}$ and $V_{\{DD\}}$. Since ϕ_i does not depend on Z_i the analysis is independent of a particular realization of observed returns.

Not all the combinations of R_i and a in Figure 6 are feasible, since $R_i \leq p - \pi_i$ and $a \leq 1/p$. Figures 7a-f depict the preference regions for the feasible area only and for various values of p and π_i. From these figures it appears that Method B performs better as R_i gets smaller, π_i gets larger or p gets smaller, particularly if $a < 1$. Note that the R_i tend to be small if the base of the time-to-return distribution, n, is large compared to the lead time L. At this time we would like to stress that the analysis of this section does not depend on time t, since equations (12)–(3) do not depend on t, nor did we make any assumptions on realizations of the demand process $\{u_i\}$ and observed returns $\{Z_i\}$.

Figure 6: Preference regions for $F_{\{V\}})$ in terms of a and R_i; $\pi_i = \beta_i = 0.6$, $p = 0.7$.



In summary, we conclude that it is not at all obvious that Method D, which is the most informed method, performs better than Method B. In fact, we have identified situations in which Method B performs better, on average, with respect to the expectation and variance of lead time net demand. In particular, when all return probabilities are underestimated or all return probabilities are overestimated Method B has opportunities to outperform Method D. In the next section we will quantify the impact of misinformation with respect to costs.

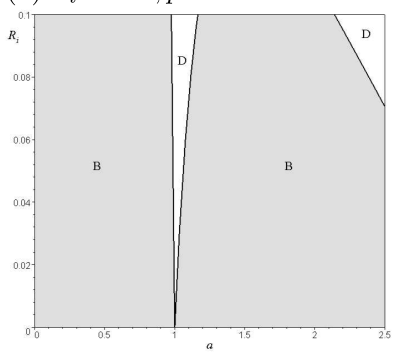
6 Cost performance

6.1 Experimental Design

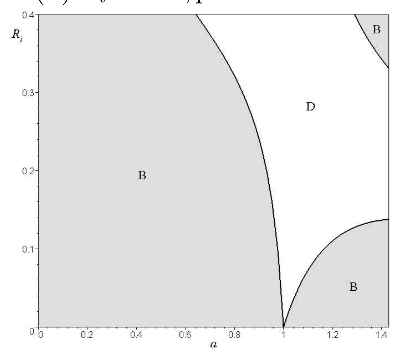
In order to quantify the impact of misinformation on inventory related cost performance we conducted a simulation study. We consider holding and backorder costs as described later. The experiments are based on the inventory system that was introduced in Section 4 and are conducted in the following manner. Each period t we draw the cumulative demand $D(t)$ from a normal distribution with mean μ_D, variance σ_D^2, and coefficient of variation $cv_D = \frac{\sigma_D}{\mu_D}$ (values are rounded to integers; negative numbers are treated as zero). For each individual item of this cumulative demand we determine the time to return based on the pre-specified return probabilities, $\{p_j\}$. Each period, estimates of the expectation and variance of the net demand during lead time are computed according to one of methods A–D. These estimates are subsequently used to compute the base stock level, S. At the end of each period, overstocks are charged with a holding cost \$ h per item, per period, whereas stockouts are penalized with \$ b per occurrence. At the end of each simulation experiment we calculate the total average cost per period as the total average holding plus backorder costs per period. Note that all methods use *estimates,*

Figure 7: Preference regions in terms of a and R_i for various values of π_i and p.

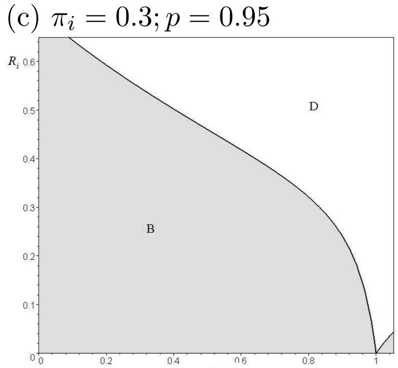
(a) $\pi_i = 0.3; p = 0.40$



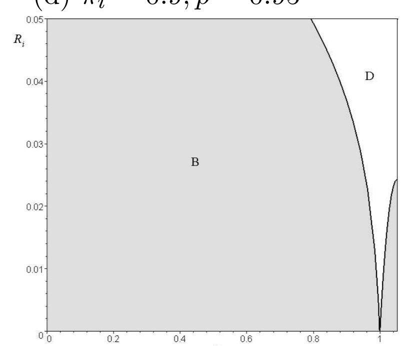
(b) $\pi_i = 0.3; p = 0.70$



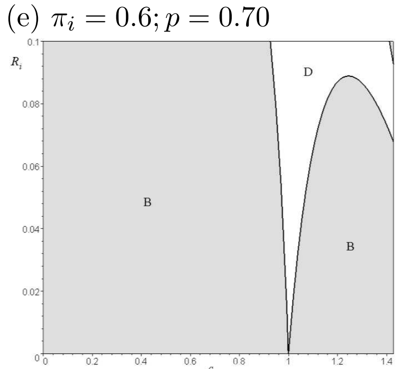
(c) $\pi_i = 0.3; p = 0.95$



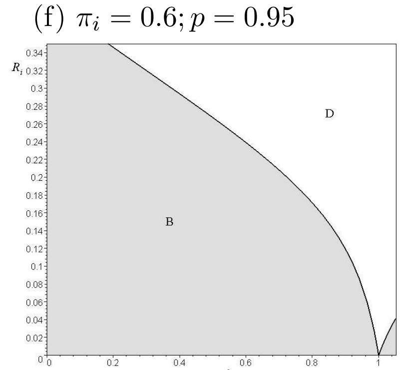
(d) $\pi_i = 0.9; p = 0.95$



(e) $\pi_i = 0.6; p = 0.70$



(f) $\pi_i = 0.6; p = 0.95$



$\{\widehat{p}_j\}$, of the real return probabilities, $\{p_j\}$, since the latter are not known. The same holds for the overall return probability, p, which is estimated as $\widehat{p}$.

Each simulation experiment consists of at least ten simulation runs of 5.000 periods, preceded by a warm-up run of the same length. The simulation stops as soon as the relative error in the total average costs is less than 1%. In order to make a better comparison among simulation experiments we make use of common random numbers. Please note that in this study all parameters are assumed to be constant over time.

Based on the estimates $\mathrm{E}[ND_L(t)]$ and $\mathrm{Var}[ND_L(t)]$ of the mean and variance of net lead time demand the base stock level S is computed as in (1). Assuming that the net demand during lead time is normally distributed, the cost optimal value of the safety factor, k, satisfies $G(k) = 1 - \frac{h}{b}$, where $G(.)$ is the standard normal distribution (see Silver et al., 1998).

The time-to-return distribution $\{p_j\}$ consists of two components: The overall return probability p and the *conditional* time-to-return probabilities $\{\overline{p}_j\}$ given that the item returns. The (unconditional) time-to-return probabilities then are defined as $p_j = p \cdot \overline{p}_j$, $j \in [0, 1, \ldots, n]$. In the simulation experiments we use two conditional time-to-return distributions. The first one is a geometric distribution with conditional expected return time $T = 1/q$, i.e. $\overline{p}_j = q(1-q)^{j-1}$, $j = 1, 2, \ldots, \infty$. The second is a discrete uniform distribution with conditional expected return time $T = (n+1)/2$, i.e. $\overline{p}_j = 1/n$, $j = 1, 2, \ldots, n$. Please note that in this study all parameters are constant in time.

6.2 Numerical study

In the case of perfect information, i.e. $\hat{p} = p$ and $\hat{p}_j = p_j$, method D will outperform all other methods since it is using all of the available information in a correct way. In order to investigate the effect of misinformation, we consider two types of errors in the parameter estimates. The first is a misspecification of the overall return probability, p, while the expected return time is preserved. The second is a misspecification of the conditional expected time-to-return, T. This affects the shape of the time-to-return distribution, but the estimated overall return probability is preserved. For example, suppose that the real time to return distribution is given by $\{p_1, p_2, p_3\} = \{0.2, 0.1, 0.2\}$. Then an estimate of $\{\widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{0.1, 0.05, 0.1\}$ would have the same conditional expected time-to-return (2.0), but a 50% lower estimated return probability. An estimate of $\{\widehat{p}_1, \widehat{p}_2, \widehat{p}_3\} = \{0.3, 0.1, 0.1\}$ would have the same overall return probability (0.5), but a lower conditional expected time-to-return.

For numerical comparisons we define the relative difference of some method m with respect to method D as follows:

$$\text{relative difference} = \frac{\text{costs method } m - \text{costs method D}}{\text{costs method D}} \times 100\%$$

6.2.1 Perfect information

Comparing the four methods in case of perfect information (Tables 6.2–6.3) we observe here that method D is indeed superior to the other methods, although the differences with respect to methods B and C are not significant for $p \leq 0.8$ (less than 1 percent). The performance of Method A is extremely poor. It uses the assumption that all lead time returns are correlated with the lead time demands. This causes a systematic underestimation of the variance in the

Table 2: Comparison of Methods A–D in case of perfect information; Geometric time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

	Methods			
	A	**B**	**C**	**D**
	relative difference*			costs
Base case: $p = 0.5, T = 1.67, cv_D = 0.2$ $L = 4, h = 1, , b = 50$	24.9%	0.2%	0.0%	26.07
$\mathbf{p = 0.8}$	111.8%	0.9%	0.4%	22.85
$\mathbf{p = 0.9}$	129.3%	3.1%	0.6%	21.24
$\mathbf{T = 2.50}$	35.7%	0.3%	0.1%	27.98
$\mathbf{T = 5.00}$	56.7%	0.4%	0.0%	30.70
$\mathbf{cv_D = 0.4}$	36.3%	0.0%	0.0%	44.80
$\mathbf{cv_D = 0.8}$	24.8%	0.0%	0.0%	83.00
$\mathbf{L = 8}$	10.6%	0.2%	0.1%	32.94
$\mathbf{L = 16}$	4.1%	0.1%	0.0%	43.37
$\mathbf{b = 10}$	11.0%	0.1%	0.2%	19.18
$\mathbf{b = 100}$	36.7%	0.1%	-0.2%	28.69

lead time net demand, especially for high return rates and large lead times. Because of this poor performance we will not consider Method A in the remainder of the numerical study.

6.2.2 Misinformation on the overall return probability

We define the relative error in the estimated return probability as follows

$$\text{relative error} = \frac{\hat{p} - p}{p} \times 100\%$$

Tables 4 & 5 show that Method B structurally outperforms the more information intensive methods C and D in case of misinformation of plus or minus 10% or more. Under an error of -20%, the relative difference can be as large as -5%. Under an error of $+20\%$, the relative difference can be as large as -20% for a return probability of 0.5 and as large as -60% for a return probability of 0.8. The performance of Method C is fairly close to the benchmark, although it usually performs worse.

The relative cost differences become bigger as the return rate goes up. This is not surprising as with increasing returns also the impact of (mis)information increases. Note that the cost improvement of Method B with respect to the benchmark can be as large as 60% under 20% overestimation and a return probability of 0.8.

An increase of the lead time makes the relative differences smaller. A longer lead time results in a larger portion of the lead time returns that come from expected issues during the lead time itself. Methods B, C, and D treat this category of returns in exactly the same way, so the relative differences become smaller.

Using a similar argument we expect that an increase in the expected time-to-return, T, results in larger relative differences. As the expected time-to-return increases the portion of lead time returns that come from expected issues during the lead time decreases. The relative differences therefore increase, as Tables 4–5 also show.

Table 3: Comparison of Methods A–D in case of perfect information; Uniform time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

	Methods A	B	C	D
	relative difference*			costs
Base case: $p = 0.5, T = 2.50, cv_D = 0.2$ $L = 4, h = 1, b = 50$	44.9%	0.3%	0.2%	28.59
$\mathbf{p = 0.8}$	171.4%	2.4%	1.1%	26.23
$\mathbf{p = 0.9}$	181.4%	4.3%	1.8%	24.83
$\mathbf{n = 4.50}$	59.6%	0.7%	0.3%	31.14
$\mathbf{n = 8.50}$	65.0%	0.3%	0.1%	32.68
$\mathbf{cv_D = 0.4}$	65.0%	0.2%	0.1%	50.22
$\mathbf{cv_D = 0.8}$	48.3%	0.2%	0.1%	93.61
$\mathbf{L = 8}$	18.2%	0.3%	0.2%	34.93
$\mathbf{L = 16}$	8.6%	0.2%	0.0%	44.80
$\mathbf{b = 10}$	24.1%	0.6%	0.2%	21.07
$\mathbf{b = 100}$	68.2%	0.6%	0.2%	31.33

Table 4: Comparison of Methods B–D in case of misestimation of the return probability; Geometric time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

	$\hat{p}$	error	Methods B	C	D
			rel. difference*		costs
Base case:	0.40	-20%	-1.9%	0.3%	35.15
$p = 0.5, T = 1.67,$	0.45	-10%	-0.9%	0.2%	29.07
$cv_D = 0.2, L = 4,$	0.55	+10%	-5.2%	0.7%	35.54
$h = 1,\ b = 50,\ k = 2.054$	0.60	+20%	-17.9%	2.4%	84.33
	0.64	-20%	-5.3%	1.2%	42.45
$\mathbf{p = 0.8}$	0.72	-10%	-3.9%	1.3%	31.18
	0.88	+10%	-43.1%	8.5%	92.26
	0.96	+20%	-67.1%	6.6%	692.04
	0.40	-20%	-3.1%	0.5%	37.10
$\mathbf{T = 2.50}$	0.45	-10%	-1.5%	0.3%	31.06
	0.55	+10%	-7.7%	1.2%	37.13
	0.60	+20%	-26.7%	4.4%	86.12
	0.40	-20%	-0.6%	0.0%	90.71
$\mathbf{cv_D = 0.8}$	0.45	-10%	-0.2%	0.0%	85.39
	0.55	+10%	-0.8%	0.0%	87.21
	0.60	+20%	-3.5%	0.2%	101.41
	0.40	-20%	-0.8%	0.1%	87.60
$\mathbf{L = 16}$	0.45	-10%	-0.7%	0.1%	61.95
	0.55	+10%	-5.1%	0.5%	170.99
	0.60	+20%	- 6.9%	0.6%	902.68
	0.40	-20%	-1.7%	0.3%	38.11
$\mathbf{b = 100}$	0.45	-10%	-0.8%	0.2%	31.79
($k = 2.326$)	0.55	+10%	-5.7%	0.8%	40.11
	0.60	+20%	-19.8%	2.9%	105.85

Table 5: Comparison of Methods B–D in case of misestimation of the return probability; Uniform time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

	$\hat{p}$	error	Methods B	C	D
			rel. difference*		costs
Base case:	0.40	-20%	-2.1%	0.3%	37.23
$p = 0.5, T = 2.50,$	0.45	-10%	-0.7%	0.3%	31.45
$cv_D = 0.2, L = 4,$	0.55	+10%	-4.9%	0.4%	36.63
$h = 1,\ b = 50,\ k = 2.054$	0.60	+20%	-18.5%	1.1%	76.85
	0.64	-20%	-4.9%	0.9%	45.24
$\mathbf{p = 0.8}$	0.72	-10%	-2.4%	1.2%	34.08
	0.88	+10%	-34.0%	-1.0%	72.54
	0.96	+20%	-56.5%	-3.8%	394.72
	0.40	-20%	-3.6%	0.6%	39.97
$\mathbf{n = 4.50}$	0.45	-10%	-1.5%	0.4%	34.21
	0.55	+10%	-7.4%	0.9%	39.01
	0.60	+20%	-28.6%	2.8%	79.98
	0.40	-20%	-0.6%	0.0%	100.03
$\mathbf{cv_D = 0.8}$	0.45	-10%	-0.1%	0.0%	95.47
	0.55	+10%	-0.4%	0.0%	97.04
	0.60	+20%	-2.8%	0.1%	107.60
	0.40	-20%	-1.0%	0.1%	88.86
$\mathbf{L = 16}$	0.45	-10%	-0.7%	0.1%	63.27
	0.55	+10%	-5.7%	0.3%	164.27
	0.60	+20%	-7.9%	0.3%	857.01
	0.40	-20%	-1.9%	0.3%	40.39
$\mathbf{b = 100}$	0.45	-10%	-0.6%	0.4%	34.40
$(k = 2.326)$	0.55	+10%	-5.3%	0.4%	41.05
	0.60	+20%	-20.5%	1.4%	93.51

An increase in demand variation also has a negative impact on the relative differences. This probably is because more demand uncertainty also leads to more return uncertainty. More information on the return distribution than has less impact on cost performance. The difference between the methods thus becomes smaller as the uncertainty in returns increases.

The analysis of Section 5 showed that in case of misinformation of the return probability Method B performs better than Method D with respect to the expected lead time net demand, while the difference between the methods increases as the return rate increases. With respect to the variance the analysis was less straightforward. Sometimes Method B outperforms Method D and sometimes it is the other way around. From the numerical results we conclude that the effect of the expected lead time net demand dominates the effect of the variance.

6.2.3 Misinformation on the conditional expected time-to-return

We define the relative error in the estimated conditional time to return as $\frac{\hat{T}-T}{T} \times 100\%$. For the geometric distribution this leads to

$$\text{relative error} = \frac{1/\hat{q} - 1/q}{1/q} \times 100\% = (\frac{q}{\hat{q}} - 1) \times 100\%$$

and for the uniform distribution we have

$$\text{relative error} = \frac{(\hat{n}+1)/2 - (n+1)/2}{(n+1)/2} \times 100\% = \frac{\hat{n} - n}{n+1} \times 100\% \;.$$

According to Tables 6 & 7, misinformation of the conditional expected time-to-return has little effect if the return probability is small. For $p = 0.8$ though, both Methods B and C perform much better than the benchmark and are far more robust with respect to misinformation (Figure 8). Again the relative differences with respect to the benchmark is positively correlated with the return probability and expected time-to-return, and negatively correlated with the lead time and demand variation.

Figure 8: Misinformation (-30%,+30%) of the expected time-to-return (Geometric time-to-return distribution, $p = 0.8, q = 0.6, \mu_D = 30, cv_D = 0.2, L = 4, h = 1, b = 50$).

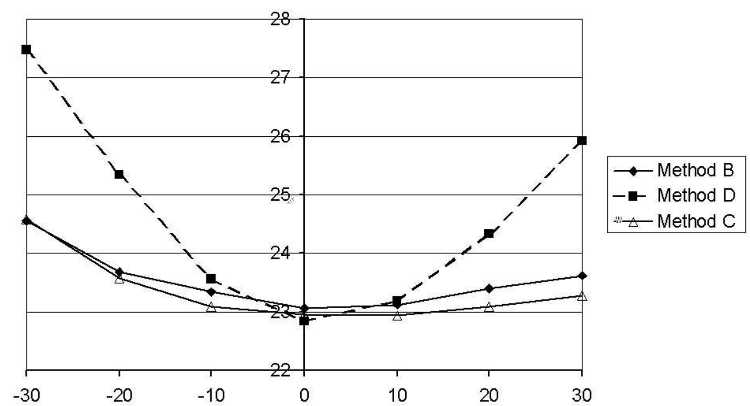


Table 6: Comparison of Methods B–D in case of misestimation of the expected time-to-return; Geometric time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

	$\hat{T}$	error	Methods B	Methods C	Methods D
			rel. difference*		costs
Base case:	1.33	-20%	-6.5%	-7.0%	25.34
$p = 0.8, T = 1.67,$	1.50	-10%	-0.9%	-2.0%	23.56
$cv_D = 0.2, L = 4,$	1.83	+10%	-0.2%	-1.1%	23.18
$h = 1,\ b = 50,\ k = 2.054$	2.00	+20%	-3.8%	-5.1%	24.33
	1.33	-20%	-0.4%	-0.6%	26.46
$\mathbf{p = 0.5}$	1.50	-10%	-0.1%	-0.3%	26.20
	1.83	+10%	0.1%	0.0%	26.05
	2.00	+20%	-0.5%	-0.2%	26.13
	2.00	-20%	-12.6%	-13.3%	30.69
$\mathbf{T = 2.50}$	2.25	-10%	-2.3%	-3.8%	26.96
	2.75	+10%	0.0%	-2.0%	26.41
	3.00	+20%	-4.8%	-6.2%	27.88
	1.33	-20%	-2.9%	-3.0%	39.95
$\mathbf{cv_D = 0.8}$	1.50	-10%	-0.4%	-0.8%	38.50
	1.83	+10%	-0.2%	-0.7%	38.30
	2.00	+20%	-2.3%	-2.9%	39.46
	1.33	-20%	-3.0%	-3.6%	32.63
$\mathbf{L = 16}$	1.50	-10%	-0.5%	-1.2%	31.65
	1.83	+10%	0.1%	-0.7%	31.35
	2.00	+20%	-3.4%	-4.3%	32.63
	1.33	-20%	-7.8%	-8.6%	28.30
$\mathbf{b = 100}$	1.50	-10%	-0.8%	-2.3%	25.86
($k = 2.326$)	1.83	+10%	0.4%	-0.8%	25.37
	2.00	+20%	-4.1%	-5.4%	26.75

Table 7: Comparison of Methods B–D in case of misestimation of the expected time-to-return; Uniform time-to-return distribution; $\mu_D = 30$. *) results below 1% are not significant

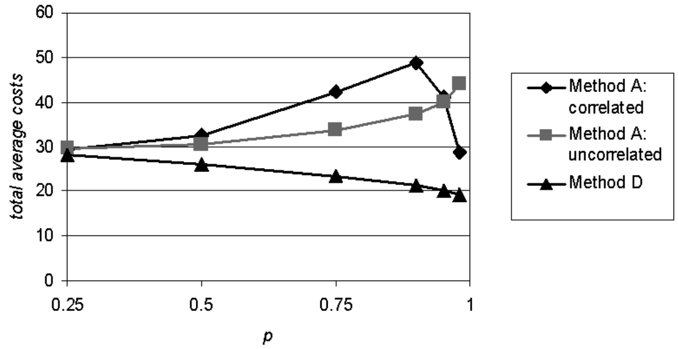
	$\hat{T}$	error	Methods **B** rel. difference*	Methods **C** rel. difference*	Methods **D** costs
Base case: $p = 0.8, T = 4.50,$ $cv_D = 0.2, L = 4,$ $h = 1,\ b = 50,\ k = 2.054$	3.50	-22%	-32.4%	5.3%	48.03
	4.00	-11%	-7.8%	2.7%	34.58
	5.00	+11%	-4.6%	-6.6%	33.40
	5.50	+22%	-16.5%	-18.0%	38.45
$\mathbf{p = 0.5}$	3.50	-22%	-3.4%	0.6%	32.75
	4.00	-11%	-0.4%	0.4%	31.57
	5.00	+11%	-0.3%	-0.7%	31.49
	5.50	+22%	-2.6%	-3.0%	32.30
$\mathbf{cv_D = 0.8}$	3.50	-22%	-12.7%	2.0%	63.38
	4.00	-11%	-1.9%	0.8%	55.29
	5.00	+11%	-3.1%	-3.9%	56.38
	5.50	+22%	-9.8%	-10.7%	61.25
$\mathbf{L = 16}$	3.50	-22%	-24.4%	4.5%	54.86
	4.00	-11%	-6.7%	2.3%	43.69
	5.00	+11%	-4.6%	-6.4%	42.66
	5.50	+22%	-16.5%	-18.4%	49.24
$\mathbf{b = 100}$ $(k = 2.326)$	3.50	-22%	-3.4%	0.7%	36.21
	4.00	-11%	-0.7%	0.4%	34.84
	5.00	+11%	-0.2%	-0.6%	34.56
	5.50	+22%	-2.6%	-2.9%	35.48

6.2.4 Extensions

Here we present a new adaptation of Method A and an extension of Method B. The disappointing performance of Method A is mainly due to the assumption that all returns during the lead time are correlated with the demands during the lead time. The other extreme, which we propose here, is to assume that all lead time returns are independent of the lead time demand. The expectation of lead time demand does not change, but the variance is then given as $\text{Var}_A[ND_L(t)] = (1+p^2)\text{Var}[D_L(t)]$. This method, denoted Method A′ - Average behavior with independence - will overestimate the variance in the lead time net demand, but it will generate less (costly) backorders (see Figure 9). Only if p is close to 1, Method A outperforms Method A′.

It was seen that Method B was rather robust under imperfect information of the time-to-return distribution. This suggests that Method B could be applied as a standard with a simple uniform distribution. The information requirement is then reduced to parameters p and n and there is no need to forecast the $\overline{p}_i$'s individually. This has significant managerial implications (as we stress below).

Figure 9: Performance of Method A both under perfect correlation and zero correlation (Geometric time-to-return distribution, $p = 0.5, q = 0.6, \mu_D = 30, cv_D = 0.2, L = 4, h = 1, b = 50$).



7 Discussion and managerial implications

Given perfect information, Method A performs in general very poorly and is not recommended for practical implementation. Thus, to *only* have knowledge on average behavior does not seem to be cost effective. Including information on the return distribution, however, does seem to provide a sufficient level of sophistication, as the performance of Method B shows. With respect to inventory related costs, the differences between Methods B, C and D are not such that they justify investments in recording detailed return data.

Both for misinformation on the return rate and on the return distribution, the differences between the methods become smaller with the decrease of the return probability and the increase of the lead time. Obviously, if there are few returns forecasting of the lead time returns is not really an issue. To understand the latter one should note that when L is large, most of the items that return during the lead time were also purchased during the lead time. The forecast of this type of returns is not based on historical data, so all Methods B–D give exactly the same forecast.

With respect to misestimating the return rate, in general it is better to underestimate the return rate than to overestimate, since stockouts are usually much more costly than overstocks. Therefore, if an interval estimate of the return rate is available, one may opt to use a value that is closer to the lowerbound rather than the upperbound.

The most robust method given misestimation of the return rate is Method B. Method B systematically outperforms Methods C and D if the return rate is misestimated by merely 10%,or more. The cost differences are particularly high if return rates are overestimated. With respect to misestimating the conditional time to return distribution, Method C and again Method B are much more robust than Method D.

The above results strongly suggest that Method B has a sufficient level of sophistication both in case of perfect estimation and imperfect estimation. In the latter case Method B is far more robust than the most informed method, Method D. For the inventory management of the mail order company of Section 2 this means that orders only have to be based on the return distribution and realized demand per period, but there is no need to track individual returns.

Finally we observed that Method B was fairly robust given misespecification of the time-to-return distribution. Therefore the following practical adaptation of the use of Method B was proposed. Companies may opt to disregard the shape of the return distribution and to simply use Method B with a flat shape, i.e. a uniform distribution. The advantages are huge in terms of information spare. The company no longer has to estimate the return probability per period, i.e. the individual $\overline{p}_i$'s. It suffices to estimate the overall return probability and the max return period, i.e. p and n. These are actual the parameters for which companies more comfortably give estimates.

8 Summary of conclusions and further research

This paper reported on the impact of (mis)information on forecasting performance and performance with respect to inventory costs by analyzing four forecasting methods as proposed by Kelle and Silver (1989). All methods make use of the expectation and variance of the demand, but different levels of information with respect to returns. The least demanding method, Method A only uses the return rate. Method B also requires the return distribution. Method C additionally uses a periodic record of returns. Finally, Method D needs to track back the period in which each individual product return was sold.

Given perfect information, forecasting performance increases as the level of information increases, Method D naturally being the best method. Yet, from the analysis we have concluded that Method B presents a reasonable level of sophistication given perfect information and is exceptionally robust given misinformation in comparison with the other methods. Furthermore since the Method B is quite insensitive to the shape of the return distribution, the companies can simply employ the uniform return distribution, i.e. using Method B with a flat shape, that is by fitting an uniform distribution.

Method D does not appear to be very robust given misestimation. This leads to the conclusion that companies are not likely to recover investments on advanced return data with inventory savings, especially in volatile environments. Naturally, it is worth to further investigate the value of return information, with respect to multiple criteria, like for instance production scheduling and human resource management (see Toktay et al., 2003 for a comparison of the methods with respect to order variability, "the bullwhip effect").

The huge gap in performance between Methods A and B suggests that refinements of Method A are possible. We proposed an enhanced Method A, i.e. Method A' was put forward, where the variance during the leadtime is underestimated. In addition, there is an opportunity to look further for similar simple methods with low information requirements but a reasonable performance.

In the analysis of misinformation, there was a static situation where a systematic error was introduced. This is natural for small errors, since it is difficult in general to reason whether observations indicate trend changes or mere stochastic behavior. However, large variations could be dealt with by an adaptive method. This is a topic for further research.

Appendix 1: Forecasting methods

Proofs can also be found in Kelle and Silver (1989).

We define

- μ_D, the expected value of Demand D
- σ_D^2, the variance of Demand D

which we use sometimes to simplify notation.

Method A - Average behavior

This method requires the following information:

- p, the overall return probability, i.e., the probability that a product is being returned eventually.

The expectation and variance of lead time net demand according to Method A are equal to

$$\mathrm{E}_A[ND_L(t)] = (1-p)\mathrm{E}[D_L(t)] \tag{4}$$

and its variance is

$$\mathrm{Var}_A[ND_L(t)] = (1-p)^2\mathrm{Var}_A[D_L(t)] + p(1-p)\mathrm{E}[D_L(t)] \tag{5}$$

Proof:

By definition, we have that,

$$\begin{aligned}
\mathrm{E}_A[ND_L(t)] &= \mathrm{E}_A[D_L(t)] - \mathrm{E}_A[R_L(t)] &(6)\\
\mathrm{Var}_A[ND_L(t)] &= \mathrm{Var}_A[D_L(t)] + \mathrm{Var}_A[R_L(t)] \\
&- 2cov[D_L(t), R_L(t)] &(7)
\end{aligned}$$

Let us first compute $\mathrm{E}_A[R_L(t)]$ and $\mathrm{Var}_A[R_L(t)]$ (Part A of the proof) and later the covariance, i.e. $cov[D_L(t), R_L(t)]$ (Part B of the proof).

Part A

$R_L(t)$ is a mixed binomial random variable, with a random number $D_L(t)$ of trials and known success probability p. Let each trial be called X_m, $m = 1, ..., D_L(t)$. Each trial has a probability of success p, independently of any other trials. Thus,

$$\begin{aligned}
\mathrm{P}[X_m = 1] &= p \\
\mathrm{P}[X_m = 0] &= 1-p
\end{aligned}$$

and

$$var(X_m) = p(1-p)$$

Since

$$R_L(t) = \sum_{m=1}^{D_L(t)} X_m$$

it follows that (see Tijms, 2003, pg. 435)

$$\begin{aligned}\mathrm{E}_A[R_L(t)] &= \mathrm{E}[X_m]\mathrm{E}[D_L(t)] \\ \mathrm{Var}_A[R_L(t)] &= \mathrm{E}^2[X_m]\mathrm{Var}[D_L(t)] + \mathrm{Var}[X_m]\mathrm{E}[D_L(t)]\end{aligned}$$

$$\mathrm{E}_A[R_L(t)] = p\mathrm{E}[D_L(t)] \tag{8}$$

$$\mathrm{Var}_A[R_L(t)] = p^2\mathrm{Var}[D_L(t)] + p(1-p)\mathrm{E}[D_L(t)] \tag{9}$$

<u>Part B</u>

Let us now compute the covariance, $cov[D_L(t), R_L(t)]$.
By definition, we have that

$$cov[D_L(t), R_L(t)] = \mathrm{E}_A[D_L(t)R_L(t)] - \mathrm{E}_A[D_L(t)]\mathrm{E}_A[R_L(t)] \tag{10}$$

with

$$\begin{aligned}\mathrm{E}_A[D_L(t)R_L(t)] &= \sum_{k=1}^{i}\sum_{i=1}^{\infty} iG_{D_L(t)} k\frac{i!}{(i-k)!(k-1)!}p^k(1-p)^{i-k} \\ &= \sum_{i=1}^{\infty} iG_{D_L(t)} \sum_{k=1}^{i} k\frac{i!}{(i-k)!(k-1)!}p^k(1-p)^{i-k} \\ &= \sum_{i=1}^{\infty}[iG_{D_L(t)}][ip] \\ &= p\sum_{i=1}^{\infty} i^2 G_{D_L(t)} \\ &= p\,[\mathrm{E}_A[D_L(t)]]^2 \\ &= p\left[\mathrm{Var}_A[D_L(t)] + [\mathrm{E}_A[D_L(t)]]^2\right]\end{aligned}$$

where we used that, E(Y)=i · p when Y is a random variable binomially distributed with parameters (i,p).
Substituting, the above and 8, in 10, we obtain the following

$$cov[D_L(t), R_L(t)] = p\,[\mathrm{Var}_A[D_L(t)]] \tag{11}$$

To finalize the proof, let us substitute 8, 9, and 11 in 6 and 7 as adequate. We get that

$$\begin{aligned}\mathrm{E}_A[ND_L(t)] &= (1-p)\mathrm{E}_A[D_L(t)] \\ \mathrm{Var}_A[ND_L(t)] &= (1-p)^2\mathrm{Var}_A[D_L(t)] + p(1-p)\mathrm{E}_A[D_L(t)]\end{aligned}$$

□

Method B - Return distribution

Suppose that we are at the end of period t. This method requires information on previous demand per period and the knowledge of the return distribution as follows:

- u_i, purchased amount during period $i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.

The expectation and variance of lead time net demand according to Method B are

$$\mathrm{E}_B[ND_L(t)] = L \cdot \mu_D - \left[\sum_{i=i_m}^{t} u_i R_i + \mu_D \sum_{i=t+1}^{t+L-1} R_i \right] \tag{12}$$

$$\begin{aligned} \mathrm{Var}_B[ND_L(t)] &= \sigma_D^2 + \sum_{i=i_m}^{t} u_i R_i (1 - R_i) \\ &+ \sum_{i=t+1}^{t+L-1} \left\{ \sigma_D^2 (1 - R_i)^2 + \mu_D R_i (1 - R_i) \right\} \end{aligned} \tag{13}$$

with $i_m = \max\{1, t - n + 1\}$ and R_i the success probability of having products from past demand u_i returning during any of the the lead time periods $t + 1, t + 2, \cdots, t + L$:

$$R_i = \begin{cases} 0, & \text{for } i < t - n + 1 \\ \sum_{j=1}^{j_m} p_{t-i+j}, & \text{for } i_m \leq i \leq t \\ \sum_{j=1}^{j_n} p_j, & \text{for } t < i < t + L \end{cases}$$

where $j_m = \min\{L, n + i - t\}$ and $j_n = \min\{n, t - i + L\}$

Proof:
Let $z_{i,j}$ be the number of products sold in period i and returned in period j, $j = 1, ..., n$. The $z_{i,j}$, $j = 1, ..., n$ are a sequence of multinomial trials with the respective distribution as follows:

$$\begin{aligned} P(z_{i,1} &= k_1, ..., z_{i,n} = k_n, z_{i,\infty} = u_i - \sum_{j=1}^{n} k_j) \\ &= \frac{u_i!}{k_1! ... k_n! (u_i - \sum_{j=1}^{n} k_j)!} p_1^{k_1} ... p_n^{k_n} p_\infty^{u_i - \sum_{j=1}^{n} k_j} \end{aligned}$$

where $z_{i,\infty}$ represents the number of products sold in period i that will never return.
We are interested in computing the total number of returns during the lead time. Let W_i be the total number of returns during the lead time periods, i.e. $t + 1, ...t + L$, coming from u_i. We have that

$$\begin{aligned} W_i &= z_{i,t+1-i} + ... + z_{i,t+L-i}, \qquad i \leqq t \\ W_i &= z_{i,1} + ... + z_{i,t+L-i}, \qquad t < i < t + L \end{aligned}$$

Being W_i a sum of multinomial trials, W_i follows a Binomial distribution, with u_i number of trials, and with a success probability given by R_i as previously defined.

Let now $\daleth_{t,L}$ be the total number of returns during the lead time periods, $t+1, ..., t+L$. $\daleth_{t,L}$ is the sum of independent W_i with $i \leq t+L-1$.

Let us distinguish between returns that come from periods until t, and returns that come from lead time demand.

<u>First part: $i \leq t$</u>

The u_i, for $i \leq t$, are observed values. Thus the expected value and variance of the correspondent total number of returns, that come back during the lead time, are as follows:

$$
\begin{aligned}
\mathrm{E}_B[R_L(t)]^{(1)} &= \sum_{i=1}^{t} \mathrm{E}[W_i] &= \sum_{i=i_m}^{t} u_i R_i \\
\mathrm{Var}_B[R_L(t)]^{(1)} &= \sum_{i=1}^{t} \mathrm{Var}[W_i] &= \sum_{i=i_m}^{t} u_i R_i (1-R_i)
\end{aligned}
$$

with $i_m = max\{1, t-n+1\}$

<u>Second part: $i > t$</u>

Here, u_i, for $i > t$, have not been observed yet. Therefore, we use $\mathrm{E}[D_L(t)] = \mu_D$ and $\mathrm{Var}[D_L(t)] = \sigma_D^2$. We obtain

$$
\begin{aligned}
\mathrm{E}_B[R_L(t)]^{(2)} &= \sum_{i=t+1}^{t+L-1} \mathrm{E}[W_i] &= \sum_{i=t+1}^{t+L-1} \mu_D R_i \\
\mathrm{Var}_B[R_L(t)]^{(2)} &= \sum_{i=t+1}^{t+L-1} \mathrm{E}[W_i] &= \sum_{i=t+1}^{t+L-1} \left[R_i^2 \sigma_D^2 + \mu_D R_i (1-R_i)\right]
\end{aligned}
$$

Finally, the expected value of the net demand during the lead time is given as follows:

$$
\begin{aligned}
\mathrm{E}_B[ND_L(t)] &= \mathrm{E}[D_L(t)]^{(2)} - \left[\mathrm{E}_B[R_L(t)]^{(1)} + \mathrm{E}[R_L(t)]^{(2)}\right] \\
&= \sum_{i=t+1}^{t+L} \mu_D - \left[\sum_{i=i_m}^{t} u_i R_i + \sum_{i=t+1}^{t+L-1} \mu_D R_i\right] \\
&= L \cdot \mu_D - \left[\sum_{i=i_m}^{t} u_i R_i + \mu_D \sum_{i=t+1}^{t+L-1} R_i\right]
\end{aligned}
$$

and the variance is given by

$$
\begin{aligned}
\mathrm{Var}_B[ND_L(t)] &= \sigma_D^2 + \sum_{i=i_m}^{t} u_i R_i (1-R_i) \\
&+ \sum_{i=t+1}^{t+L-1} \left\{\sigma_D^2 (1-R_i)^2 + \mu_D R_i (1-R_i)\right\}
\end{aligned}
$$

where equality 5 was used and with $i_m = max\{1, t-n+1\}$

□

Method C - Return distribution & return information per period

Suppose that we are at the end of period t. In addition to the requirements of method B, this method makes use of observed data on aggregated returns:

- u_i, purchased amount during period $i, i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.
- y_i, the total amount of returned products in each period $i, i \leq t$.

The expected lead time net demand according to method C is

$$\mathrm{E}_C[ND_L(t)] \quad = \quad \mathrm{E}_B[ND_L(t)] - \underline{c}\, T^{-1}(\underline{y} - E(\underline{y})) \tag{14}$$

and the variance is

$$\mathrm{Var}_C[ND_L(t)] \quad = \quad \mathrm{Var}_B[ND_L(t)] - \underline{c}\, T^{-1}\underline{c}^{\dagger} \tag{15}$$

with $\underline{y} = (y_t, y_{t-1}, ..., y_{t-n+2})$ the vector of recent aggregated returns, T the covariance matrix of vector $\underline{y}$ and T^{-1} its inverse matrix; and $\underline{c}$ a vector of covariances defined, for $j = 1, 2, \ldots, n-2$, by

$$\begin{aligned} c_j \quad &= \quad Cov\left(\textstyle\sum_{i=i_m}^{t-1} W_i\; y_{t-j+1}\right) \\ &= \quad -\textstyle\sum_{i=i_m}^{t-j} u_{t-n+i}\; p_{n-j+1-i} \sum_{m=1}^{j_m} p_{n-i+m} \end{aligned}$$

with $\underline{c}^{\dagger}$ being the transpose of vector $\underline{c}$ and $j_m = \min\{i, L\}$
And, the elements of matrix T, $T_{j,k} = cov(y_{t-j+1}, y_{t-k+1})$, are defined as follows.

$$\begin{cases} T_{j,k} = -\sum\limits_{j=i_j}^{t-k} p_{t-j+1-i}\; p_{t-k+1-i}, & \text{for } j = 1, ..., n-2,\; j \leq k \leq n-2 \\ T_{k,j} = T_{j,k}, & \text{for } k < j \\ T_{j,j} = \sum\limits_{j=i_j}^{t-k} p_{t-j+1-i}\; (1 - p_{t-j+1-i}), & \text{for } j = 1, ..., n-2 \end{cases}$$

with $i_j = \max\{1, t-j+1-n\}$

Proof:

This method makes use of the observed total amount returned in each period up to t, i.e. $y_t, y_{t-1}, ..., y_{t-n+2}$. Recent aggregated returns are correlated with the returns during the lead time. Thus, this method makes use of the conditional expectation and variance of these returns as follows:

$$\mathrm{E}_C[R_L(t)^{(1)}] = E\left[\sum_{j=i_m}^{t-1} W_i \mid y_t, y_{t-1}, ..., y_{t-n+2}\right] \tag{16}$$

$$\mathrm{Var}_C[R_L(t)^{(1)}] = Var\left[\sum_{j=i_m}^{t-1} W_i \mid y_t, y_{t-1}, ..., y_{t-n+2}\right] \tag{17}$$

The above expressions can not be expressed in an exact analytical form. However, a multidimensional normal vector gives a good approximation, if demand is reasonably large and if $n \geq 4$. The approximation allow us to write the expression 16 and 17 as follows (see Kelle and Silver, 1989):

$$\mathrm{E}_C[R_L(t)^{(1)}] = E(\sum_{j=i_m}^{t-1} W_i) + \underline{c}\, T^{-1}(\underline{y} - E(\underline{y})) \tag{18}$$

$$\mathrm{Var}_C[R_L(t)^{(1)}] = Var(\sum_{j=i_m}^{t-1} W_i) - \underline{c}\, T^{-1}\underline{c}^{\dagger} \tag{19}$$

with $\underline{y} = (y_t, y_{t-1}, ..., y_{t-n+2})$, T, and $\underline{c}$ as defined before, and T^{-1} being the inverse matrix of T, and $\underline{c}^{\dagger}$ being the transpose of vector $\underline{c}$.
Thus, the expected value and variance of the total returns during the lead time corresponds to

$$\begin{aligned}\mathrm{E}_C[ND_L(t)] &= \mathrm{E}_C[R_L(t)^{(1)}] + u_t R_t + \mathrm{E}_B[R_L(t)^{(2)}] \\ &= \mathrm{E}_B[R_L(t)] - \underline{c}\, T^{-1}(\underline{y} - E(\underline{y}))\end{aligned} \tag{20}$$

$$\begin{aligned}\mathrm{Var}_C[R_L(t)] &= \mathrm{Var}_C[R_L(t)^{(1)}] + u_t R_t(1 - R_t) + \mathrm{Var}_B[R_L(t)]^{(2)} \\ &= \mathrm{Var}_B[R_L(t)] - \underline{c}\, T^{-1}\underline{c}^{\dagger}\end{aligned} \tag{21}$$

Please note that $\mathrm{E}_C[R_L(t)]$ is as $\mathrm{E}_B[R_L(t)]$ but corrected with the term $\underline{c}\, T^{-1}(\underline{y} - E(\underline{y}))$. Similarly, $\mathrm{Var}_C[R_L(t)]$ is as $\mathrm{Var}_B[R_L(t)]$ but corrected with the term $\underline{c}\, T^{-1}\underline{c}^{\dagger}$. Therefore, we have that the expected value and the variance of the net demand during the lead time is as given next.

$$\begin{aligned}\mathrm{E}_C[ND_L(t)] &= \mathrm{E}_B[ND_L(t)] - \underline{c}\, T^{-1}(\underline{y} - E(\underline{y})) \\ \mathrm{Var}_C[ND_L(t)] &= \mathrm{Var}_B[ND_L(t)] - \underline{c}\, T^{-1}\underline{c}^{\dagger}\end{aligned}$$

□

Method D - Return distribution & tracked individual returns

Let t be the last observed period. Besides the requirements of Method B this method requires to track back in what period each individual return has been sold:

- u_i, purchased amount during period $i, i \leq t$.
- p_j, the probability of an item being returned exactly after j periods, $j = 1, ..., n$, with n being the largest j for which the return probability is non-zero.
- Z_i^t, the observed total number of product returns from each past purchase $u_i, i < t$.

To simplify notation we use next Z_i instead of Z_i^t.
The expected value and variance of the lead time net demand according to method D are respectively

$$\begin{aligned} \mathrm{E}_D[ND_L(t)] &= L \cdot \mu_D \\ &- \left[\sum_{i=i_m}^{t-1} (u_i - Z_i) Q_i + u_t R_t + \mu_D \sum_{i=t+1}^{t+L-1} R_i \right] \end{aligned} \tag{22}$$

$$\begin{aligned} \mathrm{Var}_D[ND_L(t)] &= \sigma_D^2 + \sum_{i=i_m}^{t-1} (u_i - Z_i) Q_i \left[1 - Q_i\right] + \ u_t R_t \left[1 - R_t\right] \\ &\quad + \sum_{i=t+1}^{t+L-1} \left\{ \sigma_D^2 \left[1 - R_t\right] + \mu_D R_i \left[1 - R_i\right] \right\} \end{aligned} \tag{23}$$

with $Q_i = \frac{R_i}{1 - \sum_{j=1}^{t-i} p_j}$, the success probability associated with the binomial conditional random variable W_i given Z_i, the returned amount from past demand $u_i, i \leq t$.

$$\begin{aligned} \mathrm{Var}_D[ND_L(t)] &= \sigma_D^2 + \sum_{i=i_m}^{t-1} (u_i - Z_i) Q_i \left[1 - Q_i\right] + \ u_t R_t \left[1 - R_t\right] \\ &+ \sum_{i=t+1}^{t+L-1} \left\{ \sigma_D^2 \left[1 - R_t\right] + \mu_D R_i \left[1 - R_i\right] \right\} \end{aligned} \tag{24}$$

Proof:

As before, we are interested in computing the expected value and the variance of the net demand during the lead time. For returns coming from the demand during the lead time itself, the expected value and variance of these returns is given as in Method B, i.e. by $\mathrm{E}_B[ND_L(t)^{(2)}]$ and $\mathrm{Var}_B[ND_L(t)^{(2)}]$ (also used by Method C, for these returns). For returns coming from previous demand, Method D makes use of conditional return probabilities, as it observes the returns of any particular period up to time t.
The W_i, total return during the lead time, is conditioned by Z_i, the observed total number of product returns from each u_i. We have, for $i < t$,

$$
\begin{aligned}
\mathrm{E}[W_i \mid Z_i] &= u_i - Z_i \\
\mathrm{Var}[W_i \mid Z_i] &= \frac{R_i}{1 - \sum_{j=1}^{t-i} p_j}
\end{aligned}
$$

Thus, we have that

$$
\begin{aligned}
\mathrm{E}_D[R_L(t)] &= \sum_{i=i_m}^{t-i} (U_i - V_i) Q_i + u_t R_t + \mathrm{E}_B[ND_L(t)^{(2)}] && (25) \\
\mathrm{Var}_D[R_L(t)] &= \sum_{i=i_m}^{t-1} [(U_i - V_i) Q_i (1 - Q_i)] + \\
&= u_t R_t (1 - R_t) + \mathrm{E}_B[ND_L(t)^{(2)}] && (26)
\end{aligned}
$$

with

$$
Q_i = \frac{R_i}{1 - \sum_{j=1}^{t-i} p_j}
$$

Finally, we obtain

$$
\begin{aligned}
\mathrm{E}_D[ND_L(t)] &= L \cdot \mu_D - \left[\sum_{i=i_m}^{t-1} (u_i - Z_i) Q_i + u_t R_t + \mu_D \sum_{i=t+1}^{t+L-1} R_i \right] \\
\mathrm{Var}_D[ND_L(t)] &= \sigma_D^2 + \sum_{i=i_m}^{t-1} (u_i - Z_i) Q_i \left[1 - Q_i\right] + \; u_t R_t \left[1 - R_t\right] \\
&+ \sum_{i=t+1}^{t+L-1} \left\{ \sigma_D^2 \left[1 - R_t\right] + \mu_D R_i \left[1 - R_i\right] \right\}
\end{aligned}
$$

□

Appendix 2: Forecasting Performance

In the following please note that
$\pi_i = \sum_{j=1}^{t-i} p_j$, $Q_i = R_i/(1 - \pi_i)$,
and
$\mathcal{E}_{R|D}\{Z_i\} = u_i \pi_i$.
For simplicity write '$\sum$' for '$\sum_{i=t-n+1}^{t-1}$'.

Analysis regarding the expectation of lead time net demand

Using 12, i.e.

$$\mathrm{E}_B[ND_L(t)] = L \cdot \mu_D - \left[\sum_{i=i_m}^{t} u_i R_i + \mu_D \sum_{i=t+1}^{t+L-1} R_i \right]$$

and 22, i.e.

$$\begin{aligned} \mathrm{E}_D[ND_L(t)] &= L \cdot \mu_D \\ &\quad - \left[\sum_{i=i_m}^{t-1} (u_i - Z_i)\, Q_i + u_t R_t + \mu_D \sum_{i=t+1}^{t+L-1} R_i \right] \end{aligned}$$

we have

$$\begin{aligned} E_{\{BD\}} &= \widehat{\mathrm{E}}_B[ND_L(t)] - \mathrm{E}_D[ND_L(t)] \\ &= \sum \left[(u_i - Z_i)\, Q_i - u_i \widehat{R}_i \right] && (27) \\ E_{\{DD\}} &= \widehat{\mathrm{E}}_D[ND_L(t)] - \mathrm{E}_D[ND_L(t)] \\ &= \sum (u_i - Z_i) \left(Q_i - \widehat{Q}_i \right) && (28) \end{aligned}$$

Now we analyze the cases EXP1–EXP4 as defined in section 5.1 with respect to the performance measure $F_{\{E\}}$ as defined in relation (2). Let $X \succsim Y$ denote that method X performs at least as good as method Y.

Case EXP1: $E_{\{BD\}} > 0$ and $E_{\{DD\}} > 0$

Proof:

We have that

$$\begin{aligned} F_{\{E\}} &= \sum \left[(u_i - Z_i)\, Q_i - u_i \widehat{R}_i - (u_i - Z_i)(Q_i - \widehat{Q}_i) \right] \\ &= \sum \left[(u_i - Z_i) \widehat{Q}_i - u_i \widehat{R}_i \right] \\ &= \sum \left[(u_i - Z_i) \widehat{Q}_i - u_i \widehat{R}_i \right] \\ &= \sum \left[(u_i - Z_i) \frac{\widehat{R}_i}{(1 - \widehat{\pi}_i)} - u_i \widehat{R}_i \right] \\ &= \sum \left[(u_i - u_i \pi_i) \frac{\widehat{R}_i}{(1 - \widehat{\pi}_i)} - u_i \widehat{R}_i \right] \\ &= \sum \left[(u_i \widehat{R}_i \frac{(1 - \pi)}{(1 - \widehat{\pi}_i)} - u_i \widehat{R}_i \right] && (29) \\ &= \sum \left[(u_i \widehat{R}_i \left(\frac{(1 - \pi)}{(1 - \widehat{\pi}_i)} - 1 \right) \right] && (30) \end{aligned}$$

and since $E_{\{DD\}} > 0$, follows from 28 that

$$Q_i \quad > \quad \widehat{Q}_i \tag{31}$$

because Z_i, the observed total number of product returns from past demand u_i, is by definition smaller or equal than u_i (so $u_i - Z_i$ is always larger or equal than 0).
It follows immediately from 31 that it is infeasible to consistently overestimate the return probabilities $pj's$. In the opposite case (to consistently underestimate the return probabilities $pj's$) it follows that $F_{\{E\}} \leq 0$ meaning that $B \succsim D$.

□

Case EXP2: $E_{\{BD\}} < 0$ and $E_{\{DD\}} < 0$

Proof:

With similar steps as used in 30, we have that

$$F_{\{E\}} \quad = \sum \left[(u_i \widehat{R}_i \left(1 - \tfrac{(1-\pi)}{(1-\widehat{\pi}_i)} \right) \right] \tag{32}$$

and since $E_{\{DD\}} < 0$, follows from 28 that

$$Q_i \quad < \quad \widehat{Q}_i \tag{33}$$

It follows immediately from 33 that it is infeasible to consistently underestimate the return probabilities $pj's$. In the opposite case (to consistently overestimate the return probabilities $pj's$) it follows that $F_{\{E\}} \leq 0$ meaning that $B \succsim D$.

□

Case EXP3: $E_{\{BD\}} < 0$ and $E_{\{DD\}} > 0 \Longrightarrow B \succsim D$

Proof:

We have that

$$F_{\{E\}} \quad = \quad \sum u_i \left(\widehat{R}_i (1 + \frac{(1-\pi)}{(1-\widehat{\pi}_i)}) - 2R_i \right) \tag{34}$$

and since $E_{\{DD\}} > 0$, follows from 28 that

$$Q_i \quad > \quad \widehat{Q}_i$$

It follows immediately from 35 that it is infeasible to consistently overestimate the return probabilities $pj's$. In the opposite case (to consistently underestimate the return probabilities $pj's$) it follows that $F_{\{E\}} \leq 0$ meaning that $B \succsim D$.

□

Case EXP4: $E_{\{BD\}} > 0$ and $E_{\{DD\}} < 0 \Longrightarrow B \succsim D$

Proof:

We have that

$$F_{\{E\}} \quad = \sum u_i \left(2R_i - (1 - \tfrac{(1-\pi)}{(1-\widehat{\pi}_i)})\widehat{R}_i \right) \tag{35}$$

and since $E_{\{DD\}} < 0$, follows from 28 that

$$Q_i \quad < \quad \widehat{Q}_i$$

It follows immediately from 36 that it is infeasible to consistently underestimate the return probabilities $pj's$. In the opposite case (to consistently overestimate the return probabilities $pj's$) it follows that $F_{\{E\}} \leq 0$ meaning that $B \succsim D$.

□

In summary, if we consistently overestimate or underestimate the return probabilities pj's, then $B \succsim D$.

From cases EXP1–EXP4 it is also clear that the difference between the methods increases as the R_i get bigger. In other words, for higher return rates the differences between the methods are also larger (with respect to the forecasts of the expected lead time net demand).

Acknowledgements

This research was encouraged by the Project RevLog, Proposal number ERB 4061 PL 97-0650: "Reverse Logistics and its Effects on Industry" financially supported by the European Commission. The first author is very much indebted to the financial support of the Portuguese Foundation for the Development of Science and Technology, "Fundação Para a Ciência e a Tecnologia".

References

[1] M.A. De Bodt and Van Wassenhove L.N. Cost increase due to demand uncertainty in mrp lot sizing. *Decision Sciences*, 14(3):345–353, 1990.

[2] M. P. De Brito and R. Dekker. A framework for reverse logistics. In R. Dekker, K. Inderfurth, L. van Wassenhove, and M. Fleischmann, editors, *Quantitative Approaches to Reverse Logistics*, chapter 1. Springer-Verlag, Berlin, Germany, 2003.

[3] M. P. De Brito, R. Dekker, and S. D. P. Flapper. Reverse logistics: a review of case studies. Erasmus Research Institute Report Series ERS-2003-12-LIS, Erasmus University Rotterdam, Rotterdam, the Netherlands, 2003.

[4] M. P. De Brito and E.A. van der Laan. Inventory management with returns: the impact of misinformation. Econometric Institute Report EI 2002-29, Erasmus University Rotterdam, Rotterdam, the Netherlands, 2002.

[5] R. Dekker, K. Inderfurth, L. N. van Wassenhove, and M. Fleischmann. *Quantitative Approaches to Reverse Logistics*. Springer-Verlag, Berlin, Germany, 2003.

[6] E.S. Gardner. Evaluating forecast performance in an inventory control system. *Management Science*, 36:490–499, 1990.

[7] T. N. Goh and N. Varaprasad. A statistical methodology for the analysis of the life-cycle of reusable containers. *IEE Transactions*, 18:42–47, 1986.

[8] T. Gooley. Reverse logistics: five steps to success. *Logistics Management and Distribution Report*, pages 49–54, 1998.

[9] V. D. R. Guide Jr. Product acquisition management: current industry practice and a proposed framework. *International Journal of Production Research*, 38:3779–3800, 2000.

[10] V. D. R. Guide Jr. and L.N. van Wassenhove. *Business Perspectives on Full Cycle Supply Chains*. Carnegie Mellon Press, USA, 2003.

[11] P. Kelle and E. A. Silver. Forecasting the returns of reusable containers. *Journal of Operations Management*, 8(1):17–35, 1989.

[12] A. Kokkinaki, R. Zuidwijk, J. van Nunen, and R. Dekker. Information management for reverse logistics. In R. Dekker, K. Inderfurth, L. van Wassenhove, and M. Fleischmann, editors, *Quantitative Approaches to Reverse Logistics*, chapter 16. Springer-Verlag, 2003. forthcoming.

[13] H.L. Lee, V. Padmanabhan, and S. Whang. Information distortion in a supply chain: the bullwhip effect. *Management Science*, 43(4):546–558, 1997.

[14] J. Marx-Gómez, C. Rautenstrauch, A. Nurnberger, and R. Kruse. Neuro-fuzzy approach to forecast returns of scrapped products to recycling and remanufacturing. *Knowledge-based systems*, 15(1-2):119–228, 2002.

[15] D. S. Rogers and R. S. Tibben-Lembke. *Going Backwards: reverse logistics trends and practices*. Reverse Logistics Executive Council, Pittsburgh, PA, 1999.

[16] E. A. Silver, R. Peterson, and D. F. Pyke. *Inventory management and production planning and scheduling*. John Wiley & Sons, New York, USA, 3rd edition edition, 1998.

[17] H. C. Tijms. *A first course in stochastic models*. Wiley, West Sussex, England, 2003.

[18] L. B. Toktay, E. A. van der Laan, and M. P. de Brito. Managing product returns: The role of forecasting. In R. Dekker, K. Inderfurth, L. van Wassenhove, and M. Fleischmann, editors, *Quantitative Approaches to Reverse Logistics*, chapter 3. Springer-Verlag, 2003. *forthcoming*.

[19] L. B. Toktay, L. M. Wein, and S. A. Zenios. Inventory management of remanufacturable products. *Management Science*, 46(11):1412–1426, 2000.

[20] B. Trebilcock. Return to sender. *Warehousing Management*, 9(4):24–27, 2002.

[21] E. A. Van der Laan, M. Salomon, R. Dekker, and L. N. van Wassenhove. Inventory control in hybrid systems with remanufacturing. *Management Science*, 45(5):733–747, 1999.

Publications in the Report Series Research* in Management

ERIM Research Program: "Business Processes, Logistics and Information Systems"

2003

Project Selection Directed By Intellectual Capital Scorecards
Hennie Daniels and Bram de Jonge
ERS-2003-001-LIS
http://hdl.handle.net/1765/265

Combining expert knowledge and databases for risk management
Hennie Daniels and Han van Dissel
ERS-2003-002-LIS
http://hdl.handle.net/1765/266

Recursive Approximation of the High Dimensional max Function
Ş. İI. Birbil, S.-C. Fang, J.B.G. Frenk and S. Zhang
ERS-2003-003-LIS
http://hdl.handle.net/1765/267

Auctioning Bulk Mobile Messages
S.Meij, L-F.Pau, E.van Heck
ERS-2003-006-LIS
http://hdl.handle.net/1765/274

Induction of Ordinal Decision Trees: An MCDA Approach
Jan C. Bioch, Viara Popova
ERS-2003-008-LIS
http://hdl.handle.net/1765/271

A New Dantzig-Wolfe Reformulation And Branch-And-Price Algorithm For The Capacitated Lot Sizing Problem With Set Up Times
Zeger Degraeve, Raf Jans
ERS-2003-010-LIS
http://hdl.handle.net/1765/275

Reverse Logistics – a review of case studies
Marisa P. de Brito, Rommert Dekker, Simme D.P. Flapper
ERS-2003-012-LIS
http://hdl.handle.net/1765/277

Product Return Handling: decision-making and quantitative support
Marisa P. de Brito, M. (René) B. M. de Koster
ERS-2003-013-LIS
http://hdl.handle.net/1765/278

---

* A complete overview of the ERIM Report Series Research in Management:
http://www.erim.eur.nl

ERIM Research Programs:
LIS Business Processes, Logistics and Information Systems
ORG Organizing for Performance
MKT Marketing
F&A Finance and Accounting
STR Strategy and Entrepreneurship

Managing Product Returns: The Role of Forecasting
Beril Toktay, Erwin A. van der Laan, Marisa P. de Brito
ERS-2003-023-LIS
http://hdl.handle.net/1765/316

Improved Lower Bounds For The Capacitated Lot Sizing Problem With Set Up Times
Zeger Degraeve, Raf Jans
ERS-2003-026-LIS
http://hdl.handle.net/1765/326

In Chains? Automotive Suppliers and Their Product Development Activities
Fredrik von Corswant, Finn Wynstra, Martin Wetzels
ERS-2003-027-LIS
http://hdl.handle.net/1765/363

Mathematical models for planning support
Leo G. Kroon, Rob A. Zuidwijk
ERS-2003-032-LIS
http://hdl.handle.net/1765/332

How and why communications industry suppliers get "squeezed out" now, and the next phase
L-F Pau
ERS-2003-033-LIS
http://hdl.handle.net/1765/317

Financial Markets Analysis by Probabilistic Fuzzy Modelling
Jan van den Berg, Uzay Kaymak, Willem-Max van den Bergh
ERS-2003-036-LIS
http://hdl.handle.net/1765/323

WLAN Hot Spot services for the automotive and oil industries :a business analysis or : "Refuel the car with petrol and information , both ways at the gas station "
L-F Pau, M.H.P.Oremus
ERS-2003-039-LIS
http://hdl.handle.net/1765/318

A Lotting Method for Electronic Reverse Auctions
U. Kaymak, J.P. Verkade and H.A.B. te Braake
ERS-2003-042-LIS
http://hdl.handle.net/1765/337

Supply Chain Optimisation in Animal Husbandry
J.M. Bloemhof, C.M. Smeets, J.A.E.E. van Nunen
ERS-2003-043-LIS
http://hdl.handle.net/1765/353

A Framework for Reverse Logistics
Marisa P. de Brito and Rommert Dekker
ERS-2003-045-LIS
http://hdl.handle.net/1765/354

An assessment system for rating scientific journals in the field of ergonomics and human factors
Jan Dul and Waldemar Karwowski
ERS-2003-048-LIS
http://hdl.handle.net/1765/432

Circulation of Railway Rolling Stock: A Branch-and-Price Approach
Marc Peeters and Leo Kroon
ERS-2003-055-LIS
http://hdl.handle.net/1765/902

Emerging Multiple Issue e-Auctions
Jeffrey E. Teich, Hannele Wallenius, Jyrki Wallenius and Otto R. Koppius
ERS-2003-058-LIS
http://hdl.handle.net/1765/922